a folha

Boletim da língua portuguesa nas instituições europeias

http://ec.europa.eu/translation/portuguese/magazine

N.º 45 — verão de 2014

Software livre e de código aberto — um bom amigo do tradutor

Maria José Bellino Machado
Departamento de Língua Portuguesa — Direção-Geral da Tradução — Comissão Europeia

[versão inglesa deste texto — http://ec.europa.eu/translation/portuguese/magazine/documents/folha45_foss_en.pdf]

No mundo globalizado em que vivemos, o *software* livre e de código aberto (FOSS)(1) está a tornar-se cada vez mais uma boa opção, não apenas para as administrações públicas e as empresas mas também para os utilizadores em geral.

Este artigo é um testemunho pessoal da minha experiência com *software* livre e de código aberto ao longo dos últimos cinco anos, não só como utilizadora mas também como testadora de aplicações linguísticas.

O* software *livre e de código aberto pode ser um bom amigo do tradutor?

Se ainda há alguns anos as aplicações FOSS eram difíceis de instalar e utilizar e exigiam conhecimentos informáticos substanciais, hoje em dia são simples e convivais e, de um modo geral, exigem menos recursos. Além disso, na maior parte dos casos, há muita informação — incluindo vídeos — que guia os utilizadores passo a passo.

No domínio das aplicações linguísticas, os tradutores têm agora ao seu dispor aplicações FOSS de Tradução Assistida por Computador (TAC/CAT)(2) e podem explorar o potencial da tradução automática (TA)(3) «por medida» também com *software* livre e de código aberto.

Neste artigo irei concentrar-me numa ferramenta CAT — OmegaT(4) (uma aplicação multiplataformas) — e num sistema de TA — Moses(5) (um sistema de tradução automática estatística

de vanguarda[6] que está na base de muitos serviços de TA em todo o mundo), por intermédio de um projeto pioneiro de pequena dimensão: *Moses for Mere Mortals* (Moses para os Meros Mortais)[7].

O OmegaT é uma aplicação CAT de código aberto originalmente desenvolvida por Keith Godfrey em 2000, que tem desde então sido consideravelmente melhorada e está atualmente a ser desenvolvida por uma equipa liderada por Didier Briel[8].

O OmegaT está também na base da *Suite* de Tradução Autshumato[9], no âmbito de um projeto financiado pelo Ministério das Artes e Cultura da África do Sul e desenvolvido e gerido pelo Centro de Tecnologia Textual da Universidade do Noroeste da África do Sul.

O *Moses for Mere Mortals* (MMM) nasceu em 2009 de uma iniciativa pessoal de três tradutores da DGT — João Rosas (programador) e Hilário Leal Fontes e eu própria (testadores) — que desejavam experimentar o que era então um protótipo de sistema de tradução automática praticamente só conhecido dos especialistas — Moses.

O conjunto de *scripts* MMM foi pela primeira vez disponibilizado como *software* livre de código aberto em novembro de 2009, porque gostamos de partilhar e porque pensamos que os tradutores não se devem sentir ameaçados pela tradução automática, mas sim compreender as suas vantagens e limitações e integrá-la no seu fluxo de trabalho como mais uma ferramenta CAT que os pode ajudar na sua atividade.

Além disso, os tradutores — que são afinal quem produz a matéria-prima, sob a forma de traduções humanas de elevada qualidade, que torna possível a tradução automática estatística — devem ter a possibilidade de explorar as suas memórias de tradução e utilizar a TA com meios informáticos tão acessíveis como um computador de 1000 euros ou menos... e com um dispêndio de tempo e de energia aceitável.

O facto é que, com este *software* livre e de código aberto, os tradutores podem optar por ser independentes e controlar totalmente as ferramentas da sua profissão!

Mas antes de entrar em pormenores de ordem prática, vejamos um pouco de história sobre o movimento FOSS — e a sua filosofia — e sobre o *software* livre e de código aberto e a situação atual no domínio das aplicações linguísticas.

O movimento FOSS e o sistema operativo Linux

O movimento FOSS[10] teve início em finais da década de 1970 nos EUA com Richard Matthew Stallman[11], um pioneiro e ativista que, em 1983, lançou o projeto GNU com o objetivo de criar um sistema operativo semelhante ao Unix constituído inteiramente por *software* livre.

O movimento FOSS é um vasto movimento de indivíduos que apoiam a utilização de licenças de software de código aberto. As bases da sua filosofia são: «*1 — Liberdade de utilizar o* software *para qualquer finalidade; 2 — Liberdade para alterar o* software *a fim de satisfazer as suas necessidades; 3 — Liberdade de partilhar o* software *com os seus amigos e vizinhos; 4 — Liberdade para partilhar as alterações que fizer.*»[12]

Quase duas décadas mais tarde, no início dos anos de 1990, Linus Torvalds[13], um informático finlandês (então ainda estudante), desenvolveu o embrião de um novo sistema operativo — Linux[14] —, que foi publicado pela primeira vez em 1991.

O sistema operativo Linux — nas suas diversas distribuições[15] — constitui aquilo a que se pode chamar uma notável história de sucesso e evoluiu e desenvolveu-se ao longo das duas últimas décadas com a colaboração de programadores de todo o mundo. É atualmente gerido pela Fundação Linux[16] e

Linus Torvalds continua a estar profundamente ligado ao núcleo do sistema Linux como coordenador de projeto.

O que não é do conhecimento geral é que o *software* de código aberto faz efetivamente muito mais parte da nossa vida do que poderemos pensar. Utilizamo-lo diariamente em aplicações quando acedemos à Internet ou falamos pelo telemóvel... sem nos darmos conta disso. O Linux está também a ser cada vez mais utilizado em aplicações modernas de alta tecnologia, desde frigoríficos e televisões inteligentes até automóveis de topo de gama.

Outro facto surpreendente é que 485 dos 500 supercomputadores em todo o mundo usam o sistema operativo Linux, o que representa uma esmagadora maioria de 97%(17).

Igualmente interessante é o facto de, nos últimos anos, as administrações nacionais em todo o mundo — motivadas, não só mas também, por restrições orçamentais e requisitos de confidencialidade — estarem cada vez mais a migrar para o sistema operativo Linux e para *software* livre e de código aberto em geral(18).

Nas palavras de:

NEELIE DE KROES, Vice-Presidente da Comissão Europeia e Comissária responsável pela Agenda Digital (2010-2014): «Software *de código aberto já não é um palavrão. Costumava ser um palavrão que assustava, por exemplo, em aplicações de administração pública eletrónica. O* software *de código aberto era algo que se usava em casa, como um passatempo. Ninguém ousaria jamais propor uma solução de código aberto numa organização pública. Quem dará assistência técnica? E se formos processados por violação dos direitos de propriedade intelectual de alguém se o usarmos?... Felizmente as atitudes estão a mudar!... O motivo é não só uma boa relação custo-benefício, que é fundamental na atual situação financeira, mas também uma maior liberdade de escolha. Existe uma menor dependência face aos fornecedores e os custos de mudança de fornecedor são menores. ...e as coisas estão também a mudar no setor privado. Atualmente, as grandes empresas declaram com orgulho que não só utilizam soluções de código aberto como também contribuem para elas*.»(19)

SÍTIO WEB DA ADMINISTRAÇÃO OBAMA: «Software *de Código Aberto e o Poder da Comunidade: Em projetos de* software *de código aberto, os programadores desenvolvem aplicações para resolver problemas específicos, depois disponibilizam livremente essas aplicações para que outros as possam utilizar e contribuir para a sua melhoria. A comunidade de programadores que se desenvolve em torno de projetos de* software *de código aberto de sucesso produz frequentemente aplicações que são mais seguras, flexíveis e com melhor relação custo-eficácia do que as produzidas por uma equipa a trabalhar isoladamente*.»(20)
«*Acreditamos na utilização de* software *de código aberto e na contribuição para esse mesmo* software *como uma forma de facilitar ao governo a partilha de dados, melhorar as aplicações e serviços e gerir bem o dinheiro dos contribuintes*.»(21)

PHILIP KOEHN(22), Professor Associado na Universidade de Edimburgo e Coordenador do projeto MOSESCORE cofinanciado pela UE: «*Há várias razões para criar um ambiente de investigação aberto com base na disponibilização livre dos recursos (ferramentas e* corpora*) a toda a comunidade. Uma vez que a investigação é, em grande medida, financiada por fundos públicos, é de toda a justiça que os frutos desse trabalho sejam disponibilizados ao público. O acesso livre aos recursos permite a outros grupos de investigadores desenvolverem o trabalho que já foi iniciado aqui e faculta-lhes resultados de desempenho de base com os quais podem comparar os seus próprios trabalhos inovadores.*
Embora estes objetivos sejam honrosos, a nossa motivação para criar este conjunto de aplicações é algo interesseira: a criação de sistemas de tradução automática estatística tornou-se uma tarefa extremamente complexa e os progressos rápidos neste domínio obrigam-nos a perder muito tempo a reimplementar no nosso sistema os progressos obtidos por outros investigadores. Ao reunir vários grupos de investigação para trabalharem em conjunto no mesmo sistema, há uma menor duplicação

de trabalho e podemos passar mais tempo a fazer aquilo que realmente gostamos de fazer: ter novas ideias e testá-las.»[23]

A Comissão Europeia e* software *de código aberto

A Comissão Europeia tem estado muito ativa neste domínio, promovendo o desenvolvimento e a utilização de *software* de código aberto, nomeadamente criando a Licença Pública da União Europeia (EUPL)[24] e a Iniciativa Joinup (Observatório de *Software* de Código Aberto).

Joinup é «*uma plataforma para o intercâmbio de informações, de experiências e de código baseado em* software *livre de código aberto para utilização nas administrações públicas. Esta comunidade tem como objetivo ajudar as administrações públicas a partilhar soluções informáticas, debater boas práticas e trocar experiências*»[25].

A Comissão Europeia promove também o desenvolvimento de *software* de código aberto no âmbito dos seus Programas-Quadro de Investigação e Desenvolvimento Tecnológico[26], designadamente no domínio das Tecnologias da Linguagem[27].

Foi com cofinanciamento do 7.º Programa-Quadro de IDT que o sistema de tradução automática Moses[28] foi desenvolvido e disponibilizado como *software* de código aberto e é também com financiamento deste programa-quadro que projetos como MATECAT[29] — coordenador do projeto: Marcello Federico[30] — e CASMACAT[31] — coordenador do projeto: Philipp Koehn — têm como objetivo desenvolver uma nova geração de aplicações CAT com elevados níveis de interatividade (aprendizagem em linha) e de integração entre memórias de tradução, tradução automática e bases terminológicas.

Além disso, uma vez que a existência de dados de boa qualidade é importante para as aplicações linguísticas, as instituições da UE têm disponibilizado — no portal europa.eu — *corpora* em todas as línguas oficiais da UE. Desde 2006 que o Centro Comum de Investigação da Comissão tem publicado os *corpora* alinhados da legislação da UE com mais de 10 milhões de segmentos[32], completando assim o círculo necessário para o treino de motores de tradução automática.

A Comissão Europeia apoia também a META-NET[33] — Rede da Aliança Tecnológica Europeia Multilingue —, uma vez que a disponibilidade e a integração de recursos linguísticos digitais de todos os tipos — designadamente memórias de tradução e bases terminológicas[34] — serão um fator-chave para a sobrevivência das línguas na Era Digital... e consequentemente para a sobrevivência dos tradutores!

A coleção de Livros Brancos META-NET[35] apresenta uma panorâmica da situação das línguas da UE na Era Digital. De especial interesse para Portugal, o Livro Branco *A Língua Portuguesa na Era Digital*[36] — de António Branco[37] *et al*. — deveria sensibilizar-nos para o facto de haver (imenso) trabalho a fazer!

A DGT e aplicações linguísticas de código aberto[38]

No que diz respeito à tradução automática, a DGT tem sido uma utilizadora de longa data — desde meados da década de 1970 — com um sistema comercial à base de regras. O par inglês-português foi disponibilizado em meados da década de 1980 e o par francês-português foi desenvolvido em finais da década de 1990. O Departamento de Língua Portuguesa foi um dos departamentos que utilizaram bastante a TA desde 2000 e investiu muito na sua melhoria durante quase uma década.

No entanto, o desenvolvimento e a melhoria da TA baseada em regras são dispendiosos e exigem esforços significativos, em termos de recursos humanos e de tempo, para desenvolver cada par linguístico. Em 2010, a DGT dispunha apenas de 18 combinações linguísticas (+10 como protótipos), nenhuma das quais para as novas línguas do alargamento da UE de 2004.

Por conseguinte, em 2010 foi aprovado um novo plano de ação para a criação de um novo serviço de tradução automática — MT@EC[39] — baseado no sistema de *software* de código aberto Moses. Com uma equipa liderada por Andreas Eisele[40], gestor do projeto e anteriormente investigador no projeto EuroMatrix Plus, o serviço foi disponibilizado em 2011[41].

O serviço MT@EC[42] destina-se a satisfazer as necessidades da Comissão Europeia e das instituições da UE em geral e também a ser utilizado pelas administrações nacionais da UE (incluindo as universidades)[43], oferecendo 552 combinações linguísticas entre todas as línguas oficiais da UE (62 em modo direto e as outras através de uma língua-pivô).

No que diz respeito às aplicações CAT, a DGT utilizou um produto comercial desde finais da década de 1990 e em 2012 adquiriu uma nova aplicação comercial. No entanto, a DGT tem também utilizado uma aplicação CAT de *software* de código aberto — OmegaT — para fins de prototipagem, a qual está igualmente à disposição de todos os tradutores da DGT que a queiram utilizar.

Como pode o* software *livre e de código aberto ser um bom amigo do tradutor?

No contexto brevemente descrito, qual é a situação de um tradutor — ou de um pequeno grupo de tradutores — em termos de aplicações CAT e de tradução automática?

Em termos de aplicações CAT, há uma grande variedade de aplicações comerciais e algumas aplicações de *software* de código aberto e é interessante ver o que pensam os tradutores das opções disponíveis[44].

Em termos de tradução automática, há bastante escolha entre sistemas comerciais e também serviços e plataformas que podem ser utilizados — quer gratuitamente quer a baixo custo[45].

Especificamente em termos de tradução automática de código aberto, estão disponíveis 2 sistemas principais para quem deseje desenvolver os seus próprios motores de TA e controlar plenamente todo o processo: Apertium (sistema à base de regras)[46] e Moses (de base estatística).

Aplicação CAT — OmegaT

A aplicação OmegaT é de fácil instalação e seu guia (ajuda)[47] dá explicações muito pormenorizadas. Há também vídeos na Internet que são muito úteis, tanto para começar a trabalhar com o OmegaT como para explorar as suas funções mais avançadas[48], e o seu grupo de utilizadores[49] é muito dinâmico e cooperativo.

Na DGT, o OmegaT está a ser utilizado desde 2012 para fins de prototipagem. Para o efeito, a versão OmegaT 2.6.0 foi adaptada e alargada de modo a integrar outras ferramentas da DGT. Além disso, a DGT desenvolveu internamente um «Assistente» para integrar o OmegaT no seu fluxo de trabalho (*OmegaT Project Wizard*). Em 2012, o OmegaT — e respetivo assistente — foi disponibilizado a todos os tradutores interessados.

Fiz parte do grupo de testadores tanto do OmegaT como da aplicação CAT comercial adquirida pela DGT e gostei tanto do OmegaT que o estou a utilizar há mais de dois anos. No Departamento de Língua Portuguesa, cerca de 25% dos tradutores utilizam o OmegaT como principal aplicação CAT ou em alguns projetos de tradução.

Dado que a versão do OmegaT utilizada na DGT tem algumas variantes em comparação com a versão pública, escrevi um Guia Simplificado e um Guia mais completo, a fim de ter em consideração as melhorias ou adaptações ao nosso fluxo de trabalho da versão OmegaT da DGT e do seu assistente.

Uma vez que na DGT temos o privilégio de dispor de uma Unidade de Informática que trata de todas as questões técnicas — instalação, *plugins*, *scripts*, definição de parâmetros, compatibilidade, etc. —,

estes guias incidem exclusivamente no processo de tradução. Refletem a forma como eu utilizo o OmegaT para traduzir documentos em formato Office — que são a esmagadora maioria dos documentos traduzidos na DGT — e não se destinam, de forma alguma, a cobrir todas as suas funções e possíveis utilizações.

Como gosto de partilhar — e embora algumas funções sejam específicas da DGT — ponho aqui esses guias à disposição como suplementos eletrónicos deste artigo: *DGT-OmegaT and its Wizard — Quick Guide*(50) e *DGT-OmegaT and its Wizard — A Translator's Guide*(51). Espero que possam ser úteis.

Como na DGT temos a documentação informática sobretudo em inglês, lamento, mas não fiz uma versão em português. No entanto, a documentação do OmegaT público existe também em língua portuguesa!

***Tradução automática — Moses e* Moses for Mere Mortals**

Moses é um sistema de tradução automática estatística desenvolvido com financiamento da UE no âmbito do projeto EuroMatrix(Plus) e é atualmente coordenado no âmbito do projeto MOSESCORE sob a orientação de Philipp Koehn(52).

No artigo publicado em 2011 *Haverá no futuro um milhar de sistemas de tradução automática com o Moses*?(53), Achim Ruopp(54) exprimiu um desejo: «*Adorava que florescesse um milhar de sistemas Moses e estou confiante que, com os vários cenários de utilização a tornarem-se cada vez mais fáceis, haverá muito mais de mil.*»

Em 2014, penso que é legítimo afirmar que o seu desejo se tornou realidade e que o sistema Moses está na base de muitos sistemas de tradução automática utilizados em todos os tipos de contextos.

O que é ainda mais interessante é que a tradução automática, para além de estar disponível gratuitamente na Internet — embora com algumas reservas relativamente a alguns/muitos projetos de tradução —, está também ao alcance de principiantes com computadores domésticos não sofisticados e conhecimentos informáticos básicos.

Com aplicações como *Moses for Mere Mortals* — que fazem a interface com o sistema Moses, altamente sofisticado e complexo — a tradução automática pode ser «compreensível» para utilizadores que não tenham formação académica em linguística computacional ou informática.

Quer com as suas próprias memórias de tradução quer com os *corpora* disponibilizados pelas instituições da UE — ou combinando ambos —, os tradutores podem agora utilizar esta cadeia de tradução para o mundo real, pelo menos para qualquer par de línguas oficiais da UE. Os pares linguísticos podem ser treinados num computador doméstico em poucas horas ou dias, dependendo da dimensão do *corpus* e da capacidade do computador.

As traduções geradas através do MMM — treinado com *corpora* da DGT num computador de 1000 euros — foram utilizadas no Departamento de Língua Portuguesa desde o princípio de 2010 até meados de 2011, altura em que foi disponibilizado o novo serviço MT@EC.

O MMM foi tema de um artigo n'«a folha» com um Estudo de Caso com *corpora* com o par linguístico EN-PT(55) em que é apresentado o tipo de resultados que podem ser obtidos com recursos pouco sofisticados.

O MMM é também utilizado na aplicação *MyMT* da Fundação Olanto, sediada na Suíça, como servidor de tradução de retaguarda na sua *CAT Suite*(56).

Em 2013, no âmbito do programa de tradutores visitantes da DGT(57), o meu colega Hilário Leal Fontes(58) deu formação sobre *Moses for Mere Mortals* a estudantes e professores das universidades do

Porto[59] e do Minho. Tanto quanto sei, Moses — através dos *scripts* MMM — foi treinado diretamente e utilizado com *corpora* específicos em duas teses de mestrado juntamente com outros sistemas de tradução automática disponíveis publicamente.

Isto demonstra que é possível dispor de uma variedade «caseira» de tradução automática… e que querer é poder!

Como contribuir para o movimento FOSS

É interessante notar que — no domínio das aplicações linguísticas FOSS — os europeus têm sido líderes e dado grandes contributos, apoiando e financiando projetos de investigação ou desenvolvendo, por iniciativa privada, aplicações de código aberto. Mas é claro que as aplicações FOSS não têm fronteiras e há contribuições de todo o mundo.

O facto é que a utilização de aplicações FOSS não é meramente uma questão de poupar dinheiro utilizando *software* «baratucho»; trata-se, isso sim, de partilhar o melhor *software* possível desenvolvido em cooperação aberta.

Cada um de nós pode contribuir de várias formas: passando a palavra, utilizando esse *software*, contribuindo com a sua experiência — quer se trate de um vídeo, de um guia, de ajudar os outros ou de fazer sugestões para melhorias —, partilhando código ou fazendo um donativo.

Não há nada na vida que possa ser considerado um dado adquirido e o *software* livre e de código aberto não é uma exceção. O movimento FOSS precisa de ser cultivado e apoiado para sobreviver e a Internet — e todos nós — torna isso possível!

maria.machado@ec.europa.eu

---

(1) Wikipedia, *Free and open-source software*, http://en.wikipedia.org/wiki/Free_and_open-source_software.
(2) Wikipedia, *Computer-assisted translation*, http://en.wikipedia.org/wiki/Computer-assisted_translation.
(3) Wikipedia, *Machine translation*, http://en.wikipedia.org/wiki/Machine_translation.
(4) Wikipedia, *OmegaT*, http://en.wikipedia.org/wiki/OmegaT; OmegaT, http://www.omegat.org/.
(5) Wikipedia, *Moses (machine translation)*, http://en.wikipedia.org/wiki/Moses_(machine_translation);
Statistical Machine Translation, *Moses*, http://www.statmt.org/moses/.
(6) YouTube, *Philipp Koehn, Daniel Marcu, Kevin Knight, William Wong: European Inventor Award 2013 nominees in the research category. Inventors*, EPOFilms, https://www.youtube.com/watch?v=txMe9-geopk.
(7) Wikipedia, *Moses for Mere Mortals*, http://en.wikipedia.org/wiki/Moses_for_Mere_Mortals;
Google Code, *moses-for-mere-mortals*, http://code.google.com/p/moses-for-mere-mortals/.
(8) Mironov, R., *OmegaT Heroes: Didier Briel – Part 1; Part 2*, 21.2.2014,
http://www.velior.ru/2014/02/21/meet-omegats-development-manager/.
(9) Autshumato, http://autshumato.sourceforge.net/.
(10) Wikipedia, *Open-source movement*, http://en.wikipedia.org/wiki/Open-source_movement.
(11) Wikipedia, *Richard Stallman*, http://en.wikipedia.org/wiki/Richard_Stallman.
(12) YouTube, *The Story of Linux*, The Linux Foundation,
https://www.youtube.com/watch?v=5ocq6_3-nEw&list=PLt6G8wJalAaTmjY19Wj_SfJ2HiJgRzDLl&index=27.
(13) Wikipedia, *Linus Torvalds*, http://en.wikipedia.org/wiki/Linus_Torvalds;
YouTube, *Linus Torvalds, Linux and GNU/Linux*, AI's Perspectives and Technology,
https://www.youtube.com/watch?v=kZlOCHYu1Vk&list=PLt6G8wJalAaTmjY19Wj_SfJ2HiJgRzDLl.
(14) Wikipedia, *Linux*, http://en.wikipedia.org/wiki/Linux.
(15) Wikipedia, *Linux distribution*, http://en.wikipedia.org/wiki/Linux_distribution.
(16) Wikipedia, *Linux Foundation*, http://en.wikipedia.org/wiki/Linux_Foundation.
(17) Top 500 Supercomputer Sites, *Statistics*, http://www.top500.org/statistics, consultado em 24.8.2014.
(18) Wikipedia, *List of Linux adopters*, http://en.wikipedia.org/wiki/List_of_Linux_adopters.
(19) YouTube, *Neelie Kroes on open source and the importance of communities*, GUADEC,
https://www.youtube.com/watch?v=ok100U4Fo3Y.
(20) Welsch, P., *Open Source and the Power of Community*, The White House Blog, 5.12.2012,
http://www.whitehouse.gov/blog/2012/12/05/open-source-and-power-community.
(21) The White House, *Developers*, http://www.whitehouse.gov/developers.

[22] Wikipedia, *Philipp Koehn*, http://en.wikipedia.org/wiki/Philipp_Koehn;
Philipp Koehn, http://homepages.inf.ed.ac.uk/pkoehn/.
[23] Koehn, P., *et. al.*, *Final Report of the 2006 Language Engineering Workshop — Open Source Toolkit for Statistical Machine Translation: Factored Translation Models and Confusion Network Decoding*, https://ufal.mff.cuni.cz/~bojar/publications/2006-FILE-koehn_etal_jhuws_2006-2006-jhu-report.pdf.
[24] Wikipedia, *European Union Public Licence*, http://en.wikipedia.org/wiki/European_Union_Public_Licence.
[25] Joinup, *Open source observatory*, https://joinup.ec.europa.eu/community/osor/home.
[26] European Commission, *Horizon 2020: The EU Framework Programme for Research and Innovation*, http://ec.europa.eu/programmes/horizon2020/;
Europa, *Sínteses da legislação da UE — Sétimo Programa-Quadro (2007-2013)*, http://europa.eu/legislation_summaries/energy/european_energy_policy/i23022_pt.htm.
[27] Community Research and Development Information Service (Cordis), *Language Technologies*, http://cordis.europa.eu/fp7/ict/language-technologies/.
[28] The EuroMatrix Project (Sept. 2006 - Febr. 2009), http://www.euromatrix.net/;
EuroMatrixPlus, 2009-2012, http://www.euromatrixplus.net/;
Statistical Machine Translation, *MosesCore*, 2012-2014, http://www.statmt.org/mosescore/.
[29] MateCat, http://www.matecat.com/;
YouTube, *MateCat Project: a simple, web-based CAT Tool for translators*, Fondazione Bruno Kessler, https://www.youtube.com/watch?v=ZGQesDmoF6c;
YouTube, *MateCat — Getting Started*, MateCat, https://www.youtube.com/watch?v=LQSnBQWEZeU (com uma demonstração utilizando o português como língua de chegada!).
[30] Marcello Federico, https://sites.google.com/site/marcellofedericohome/.
[31] CASMACAT, http://www.casmacat.eu/;
YouTube, *CasMaCat Press Release UPV* [Universitat Politècnica de València], prhltupv, https://www.youtube.com/watch?v=cC9jYiNmSFQ.
[32] European Commission, *Joint Research Centre - Language Technology Resources*, http://ipsc.jrc.ec.europa.eu/?id=61.
[33] META-NET, http://www.meta-net.eu/
[34] Correia, P., Fontes, H. L., «Interações práticas entre corpora e terminologia na DGT — o papel das memórias de tradução», *in* «a folha», n.º 42 — verão de 2008, http://ec.europa.eu/translation/portuguese/magazine/documents/folha42_pt.pdf.
[35] META-NET, *META-NET White Paper Series*, http://www.meta-net.eu/%20whitepapers/index_html/.
[36] Rehm, G. (org.), Uszkoreit, H. (org.), *A Língua Portuguesa na Era Digital*, META-NET, Coleção Livros Brancos, Springer, http://www.meta-net.eu/whitepapers/e-book/portuguese.pdf.
[37] António Horta Branco, http://www.di.fc.ul.pt/~ahb/.
[38] Directorate-General for Translation, *Translation Tools and Workflow*, EU Bookshop, http://bookshop.europa.eu/is-bin/INTERSHOP.enfinity/WFS/EU-Bookshop-Site/en_GB/-/EUR/ViewPublication-Start?PublicationKey=HC3212080.
[39] Interoperability Solutions for European Public Administrations (ISA), *Machine Translation Service*, http://ec.europa.eu/isa/actions/02-interoperability-architecture/2-8action_en.htm.
[40] Andreas Eisele, http://www.coli.uni-saarland.de/~eisele/.
[41] Eisele, A., Lavecchia, C., «Using Statistical Machine Translation for Computer-Aided Translation at the European Commission», *Proceedings of the Third Joint EM+/CNGL Workshop "Bringing MT to the User: Research Meets Translators (JEC'11)*, Luxemburgo, 14.10.2011, http://mt-archive.info/JEC-2011-Eisele.pdf.
[42] Directorate-General for Translation, «Machine Translation», *Languages and Translation*, n. 6, 02/2013, http://ec.europa.eu/dgs/translation/publications/magazines/languagestranslation/documents/issue_06_en.pdf.
[43] Directorate-General for Translation, *MT@EC: European Commission machine translation for Public Administrations in the EU Member States*, http://ec.europa.eu/isa/documents/presentations/european-commission-machine-translation-for-public-administrations-in-the-eu-member-states_en.pdf.
[44] ProZ.com: The translation workplace, *Software Comparison Tool: CAT Tools*, http://www.proz.com/software-comparison-tool/cat/cat_tools/2.
[45] Wikipedia, *Comparison of machine translation applications*, http://en.wikipedia.org/wiki/Comparison_of_machine_translation_applications.
[46] Wikipedia, *Apertium*, http://en.wikipedia.org/wiki/Apertium;
Apertium: http://www.apertium.org/index.eng.html#translation.
[47] Welsh, S., Prior, M., *OmegaT for CAT Beginners*, http://www.omegat.org/en/tutorial/OmegaT%20for%20Beginners.pdf.
Smolej, V., *OmegaT 3.0 User's Guide*, http://ob.nubati.net/ditundat/omegat/docs/en30/index.html.
[48] Por exemplo, os vídeos de Roman Mironov no YouTube, https://www.youtube.com/user/veliortranslations/videos.
[49] OmegatT User Support, https://groups.yahoo.com/neo/groups/omegat/info.
[50] Machado, M. J., *DGT-OmegaT and its Wizard — Quick Guide*, http://ec.europa.eu/translation/portuguese/magazine/documents/folha45_omega-t_quick_guide_en.pdf.
[51] Machado, M. J., *DGT-OmegaT and its Wizard — A Translator's Guide*, http://ec.europa.eu/translation/portuguese/magazine/documents/folha45_omega-t_translator_guide_en.pdf.
[52] Machado, M. J., Fontes, H. L., Rosas, J., «Sistema de tradução automática Moses e... *Moses for Mere Mortals*», *in* «a folha», n.º 32 — primavera de 2010, http://ec.europa.eu/translation/portuguese/magazine/documents/folha32_pt.pdf.
[53] Ruopp, A., «Will there be a thousand Moses MT systems?», *TAUS: Enabling Better Translation*, 16.2.2011, https://www.taus.net/articles/will-there-be-a-thousand-moses-mt-systems.

[54] Achim Ruopp, http://langtechnews.hivefire.com/topic/profile/achim-ruopp/.
[55] Machado, M. J., Fontes, H. L., «Machine Translation: Case Study — English into Portuguese — Evaluation of Moses in DGT Portuguese Language Department using Moses for Mere Mortals» *in* «a folha», n.º 37 — outono de 2011, http://ec.europa.eu/translation/portuguese/magazine/documents/folha37_moses_en.pdf.
[56] Olanto Foundation, http://olanto.org;
Olanto, *myMT v2.0: Installation Manual for the Translation (Back-End) Server*, http://olanto.org/docs/Installing_Back-End_Server_V2.0.pdf.
[57] Directorate-General for Translation, *Visiting translator scheme*, http://ec.europa.eu/dgs/translation/programmes/visiting/index_en.htm
[58] I Conferência Internacional de Tradução e Tecnologia, 13 e 14 de maio de 2013, http://conferenciatraducao.wix.com/traducao-automatica#;
Fontes, H. L., *Tradução Automática para tradutores (e não só) na Comissão Europeia*, I Conferência Internacional de Tradução e Tecnologia, 13 e 14 de maio de 2013, http://www.slideshare.net/conferenciatraducao/hilrio-fontes-traduo-automtica-na-ce-22062540.
[59] Maia, B., *Revising machine translation — a marketable skill*, Colóquio Internacional LIDILE EA 3874 – Projeto Europeu OPTIMALE, 6 de junho de 2013, http://www.translator-training.eu/attachments/article/65/Revising%20MT.ppt.

Corpora multilingues da União Europeia para reutilizar na tradução

Hilário Leal Fontes
Direção-Geral da Tradução — Comissão Europeia

[versão inglesa deste texto — http://ec.europa.eu/translation/portuguese/magazine/documents/folha45_corpora_en.pdf]

Os *corpora* bi- ou multilingues paralelos estão na base do que hoje melhor se faz em termos de tradução automática estatística. São também imprescindíveis para encontrar terminologia em contexto. Para termos uma ideia das possibilidades em matéria de tradução automática com base em *corpora* paralelos, basta usar motores de tradução disponíveis gratuitamente na Web, como o Google Translate[1] ou o Bing Translator[2]. Para a terminologia em contexto, os serviços de concordanciação como os prestados por Linguee[3], Glosbe[4] ou MyMemory[5] são bons exemplos do que é possível fazer. Embora hoje em dia incorporem muitas outras fontes bi- ou multilingues, estes serviços e ferramentas desenvolveram-se a partir dos *corpora* paralelos da União Europeia (UE) e, para línguas como o chinês, o árabe e o russo, a partir dos *corpora* paralelos das Nações Unidas.

No presente artigo gostaria de dar uma breve panorâmica dos *corpora* paralelos construídos a partir dos textos multilingues da União Europeia que podem ser descarregados e utilizados diretamente com ferramentas de tradução assistida por computador e/ou usados para a construção de motores de tradução automática estatística. Antes disso, porém, recuemos um pouco para se ter uma ideia de onde vêm e do potencial inexplorado que os textos publicados pela UE ainda encerram.

A UE traduz milhões de páginas por ano. Só a Direção-Geral da Tradução (DGT) da Comissão Europeia traduz atualmente cerca de dois milhões de páginas por ano[6], incluindo cerca de 80 mil para português. Se considerarmos que a Comissão emprega um pouco mais de um terço dos tradutores da UE, poderemos pensar que a UE no seu conjunto produzirá mais de 200 mil páginas por ano de traduções de outras línguas para português. Se é verdade que uma parte destas páginas não se destina a publicação (comunicação com o Estado português, cartas aos cidadãos e/ou às organizações que contactam a Comissão em português, etc.), a grande maioria das páginas produzidas pela Comissão em português acaba por ser publicada nas séries C e L do *Jornal Oficial da União Europeia* (JOUE)[7], na EU Bookshop[8] e no portal europa.eu[9]. A esmagadora maioria das páginas produzidas pela DGT para português é traduzida igualmente para todas ou quase todas as outras línguas oficiais da UE. Não me parece que a situação nos outros serviços de tradução seja fundamentalmente diferente[10].

Tendo em vista a reutilização das traduções, na DGT as traduções e as publicações são alinhadas à frase e armazenadas numa base de dados chamada Euramis, sendo a partir desta base reutilizadas com os programas de tradução assistida ao nosso dispor. Esta grande memória constitui igualmente a base para construir os motores do serviço de tradução automática da Comissão, o MT@EC. A base Euramis não está disponível ao público, mas a DGT publica com regularidade o conteúdo da série L (legislação) do JOUE alinhado à frase, de forma a que tradutores e investigadores e todos aqueles que vejam utilidade nos *corpora* paralelos alinhados deles possam tirar proveito. A DGT não esteve sozinha neste tipo de ação. Na verdade, os pioneiros a publicar *corpora* paralelos alinhados dos textos da UE em grande escala foram, a título individual, o prof. Philipp Koehn[(11)], com a publicação do *corpus* Europarl, e o Centro Comum de Investigação da Comissão (JRC na sigla inglesa), que publicou o acervo comunitário — toda a legislação da UE em vigor num dado momento — em todas as línguas, incluindo as das adesões de 2004.

Vejamos então em seguida os *corpora* paralelos alinhados com textos da UE que estão disponíveis para o público.

Europarl[(12)]

É provavelmente o *corpus* paralelo multilingue mais utilizado na investigação e na construção de motores de tradução automática com línguas europeias. As sucessivas versões deste *corpus* foram acrescentando os textos mais recentes aos mais antigos e corrigindo um ou outro problema técnico. São constituídas pela transcrição dos debates do Parlamento Europeu e foram compiladas pelo prof. Philipp Koehn, que necessitava de um *corpus* multilingue para a investigação no domínio da tradução automática estatística, nomeadamente para o desenvolvimento do projeto Moses. Atualmente, vai na versão 7 e contém mais de 2 milhões de frases alinhadas entre o inglês e o português. A linguagem é bastante colorida, típica dos debates parlamentares.

Acervo comunitário, versão JRC[(13)]

O JRC publicou em 2006 o *corpus* do «acervo comunitário» em 20 línguas oficiais, com cerca de 9 milhões de palavras por língua. A última versão publicada, de janeiro de 2014, abrange 22 línguas e contém, em média, quase 48 milhões de palavras por língua.

Acervo comunitário, versão DGT[(14)]

Em rigor, a designação «acquis» (acervo comunitário) não é correta. Se o projeto inicial do JRC era compilar a legislação em vigor em 1 de maio de 2004 (era esta que seria traduzida nas línguas do alargamento de 2004), o projeto da DGT resultou da compilação das séries C (Comunicações e informações) e L (Legislação) do JOUE publicadas a partir de maio de 2004. Atualmente abrange os textos publicados até ao final de 2011. O alinhamento aqui está feito ao parágrafo, sendo possível obter, a pedido, os ficheiros XML de base. Contém quase 5 milhões de parágrafos por língua e abrange 23 línguas (ainda não inclui o croata). A sua utilização diretamente em ferramentas de tradução assistida ou na construção de motores de tradução automática sem antes fazer uma segmentação mais fina poderá não produzir os melhores resultados. Em termos de conteúdo há uma sobreposição com a memória da DGT apresentada a seguir, na medida em que contém a série L, mas a série C é inédita e a linguagem, sem deixar de ser administrativa, menos legislativa. Ainda este ano deverá ser publicada uma nova versão incluindo os textos mais recentes. Esta coleção constitui uma excelente base para uma universidade (alô Minho!) alinhar e publicar com uma granularidade mais fina.

Memória de tradução da DGT[(15)]

A Direção-Geral da Tradução da Comissão Europeia publicou em 2007 uma coleção de documentos legislativos alinhados e verificados manualmente nos seus diferentes departamentos linguísticos ao longo dos dez anos anteriores. Em 2011 foi publicada uma nova coleção com o alinhamento

automático e não verificado manualmente da série L do *Jornal Oficial da União Europeia* a partir de maio de 2004 até ao fim de 2010. Em 2012 e 2013 teve incrementos anuais correspondentes aos atos publicados nos anos anteriores. Atualmente a coleção está publicada em 22 línguas (todas as línguas oficiais exceto o irlandês e o croata) e podem extrair-se *corpora* bilingues em qualquer par das línguas que o constituem. A linguagem é, em parte, a linguagem típica dos atos legislativos, mas muitos dos textos têm conteúdos técnicos.

Outras coleções do JRC

O JRC publicou também pequenas memórias de tradução com textos da DG Educação e Cultura (EAC)(16) e do Centro Europeu de Prevenção e Controlo das Doenças(17).

Projeto Per-Fide(18)

Trata-se de uma iniciativa muito louvável da Universidade do Minho, que tratou com as suas ferramentas várias versões dos *corpora* publicados pela DGT e pelo JRC em seis línguas (DE, EN, ES, FR, IT e PT), compilou uma coleção inédita de textos do Banco Central Europeu e compilou ainda vários outros *corpora* paralelos que não têm a ver diretamente com a UE. Dentro do mesmo projeto, há também uma interface de concordanciação(19) que permite pesquisar nos *corpora* já referidos e ainda noutros só disponíveis para consulta em linha.

Opus(20)

Este sítio contém toda uma série de coleções de *corpora* paralelos alinhados com um sem-número de línguas. No que respeita aos *corpora* com origem em textos da UE e que contêm português, encontram-se os *corpora* paralelos alinhados seguintes:

a) *EMEA*, textos da Agência Europeia de Medicamentos; b) *Banco Central Europeu* (o *corpus* compilado pela Universidade do Minho); c) *Europarl*, versões 3 e 7; d) *Constituição Europeia*; e) *EUBookshop*. Este último *corpus* resulta de uma compilação dos textos publicados nesta livraria virtual da UE.

hilario.fontes@ec.europa.eu

---

(1) Google Tradutor, https://translate.google.com/.
(2) Bing Tradutor, http://www.bing.com/translator/.
(3) Linguee, http://www.linguee.com/.
(4) Glosbe, http://glosbe.com/
(5) MyMemory, http://mymemory.translated.net/.
(6) 1 página = 1500 carateres sem espaços.
(7) EUR-Lex, *Jornal Oficial da União Europeia*, http://eur-lex.europa.eu/oj/direct-access.html?locale=pt.
(8) EU Bookshop, https://bookshop.europa.eu/pt/home/.
(9) Europa, http://europa.eu/index_pt.htm.
(10) As traduções produzidas pelo serviço de tradução do Tribunal de Justiça são publicadas na Coletânea da Jurisprudência, http://bookshop.europa.eu/pt/bundles/colet-neas-da-jurisprud-ncia-do-tribunal-de-justi-a-do-tribunal-geral-e-do-tribunal-da-fun-o-p-blica-cbbpqep2IxiOAAAAE1GKQ16Jon/ e na base EUR-Lex (*Legislação da UE e documentos conexos - Jurisprudência da UE*), http://eur-lex.europa.eu/collection/eu-law/eu-case-law.html.
(11) Philipp Koehn, http://homepages.inf.ed.ac.uk/pkoehn/.
(12) Statistical Machine Translation, *European Parliament Proceedings Parallel Corpus 1996-2011*, http://www.statmt.org/europarl/.
(13) Joint Research Centre, *Language Technology Resources: JRC-Acquis*, http://ipsc.jrc.ec.europa.eu/index.php?id=198.
(14) Joint Research Centre, *Language Technology Resources: DGT-Acquis*, http://ipsc.jrc.ec.europa.eu/index.php?id=783.
(15) Esta coleção inicial deixou de estar disponível em linha em 2013.
(16) Joint Research Centre, *Language Technology Resources: EAC-TM*, http://ipsc.jrc.ec.europa.eu/index.php?id=784.
(17) Joint Research Centre, *Language Technology Resources: ECDC-TM*, http://ipsc.jrc.ec.europa.eu/index.php?id=782.
(18) Per-Fide, *Recursos*, http://per-fide.ilch.uminho.pt/site.pl/resources.pt.
(19) Per-Fide, *Query: Type: bilingual + Select language: PT-EN*, http://per-fide.ilch.uminho.pt/query/bilingual/PT-EN.
(20) Corpus, http://opus.lingfil.uu.se/.

Baía de Cascais

Luís Filipe PL Sabino
Antigo funcionário — Comissão Europeia; Comité Económico e Social Europeu-Comité das Regiões

Outubro

Em outubro de 1986 estava em Reykjavík. Colateralmente, jornalistava eu então, além do mais, para umas gazetas sobre coisas da guerra e do mundo em geral e sobre preocupações da humanidade, que a tanto muito propenso sou. Antes passara por Hamburgo para um «seminário» sobre Rudi Dutschke, que conhecera em 1969, e sobre quem escrevera uns textos, posteriormente, a solicitação de uma revista estudantil. Coisa pouca mas excitante, ao que parece.

Já antes havia estado na Islândia(1); agora não havia hotéis livres e o dia enfarruscara-se(2) no frio ártico. A cimeira Reagan-Gorbachev, em 11 e 12 de outubro(3), trouxera a Reykjavík paletes de políticos, jornalistas, TVs, secretas, KGBs, CIAs, conselheiros de todo o quilate, enfim gente do mundo de então de um lado e do outro da Cortina de Ferro prestes a ruir(4). Conseguira, mesmo assim, alojar-me perto da Höfði(5), em casa de um colega islandês dos meus tempos na polícia, com quem trabalhara na investigação da morte suspeita de um cidadão islandês ligado a Portugal, em tempos diplomata num país continental, cujo corpo flutuava preso a um batelão, às 06h05 num canal perto de Gent (Bélgica). Corria, então, o ano do Senhor de 197...

Vera

A Vera K. nascera em Leninegrado. Tinha cabelo cor de lava (apropriada para o cenário islandês) e a pele branca quase tipo glaciar Vatnajökull, e suave à maneira do sabonete Lux das 9 de cada 10 estrelas. 1m e 75. Falava além do russo, o inglês, o português do Brasil — estivera na embaixada soviética neste país e em Angola em 1975 —, o castelhano, o francês, o alemão e o italiano. O pai, herói soviético estalinegradista género stakhanovista, teria sido detido por um eventual acidente de viação e nunca mais reapareceu em circulação (nem eventual viatura, nem corpo nem nada), vai-se lá saber porquê. Ela apesar de tudo acreditava naquilo a que chamava socialismo. Com uma certa ataraxia, afirmava não lamentar nada(6) do que fizera pelo poder soviético. Acompanhava agora a delegação soviética com Gorbachev, afirmando ser correspondente não sei de quê e ao que tal coisa correspondia. Era claramente do aparelho e mais do que isso. Já a encontrara em tempos no Cairo, quando os soviéticos se lembraram de invadir o Afeganistão em AD 1979. No bar do hotel disse-me que estava ali, no Egito, em "serviço" (sem explicar mais) e que de Portugal só se lembrava da baía de Cascais... o que não era mau para começar a conversa(7). Uma baía dá para tudo. E que em Lisboa comprara um disco do Manuel de Falla com as *Noches en los jardines de España*.

Reykjavík

Fomos jantar para os lados da Ingólfsstraeti. Falámos em russo. Que se divorciara do marido alemão da DDR, oficial superior na VOPO; que este acabara assassinado na Bulgária, numa estância de verão do Partido — causas? Desconhecidas; que não tinha filhos e que não acreditava que a Cortina desabasse um dia. Aliás, dizia, a situação atual tornava o mundo mais seguro, um mundo bipolar, sendo melhor coabitar em condomínio fechado do que em moradias dispersas... Que após o Vietname, com que os EUA nada aprenderam(8), o Ocidente estava em declínio acentuado. Argumentei que não, que os regimes da URSS e satélites aluiriam(9) e que isso estava à vista, à luz do desastre moral e económico. Que aquela cimeira ali em Reykjavík conduziria a isso ou era um passo mais para isso. Mas eu compreendia a posição em que ela se encontrava. E, surpreso, passou-me inesperadamente *infos* que não percebi se era desinformação ou algo de sério que queria dizer, mas que receava, até porque sabia que no dia seguinte eu ia a reunião com John Poindexter e outros *top* à mistura com *star wars*, START, NORAD, NATO, etc.

14 de outubro de 1986
Zil

Num bar da capital islandesa. À porta um Zil preto, o motorista na noite gelada, *ushanka* com estrela, foice e martelo, ar de armário antigo, encostado ao Zil, tipo motorista de Ministério(10). Lá dentro, o coronel Anatoly T., que conheci em receção de embaixada meses antes e que sabia pertencer ao GRU (*military intelligence*), veio ter comigo embrulhado em rum e charuto cubano, ar falsamente contristado: que a Vera K. sofrera um infeliz acidente de viação na noite anterior e que não a tinham podido salvar... Telefonei ao Sigurdur, da polícia islandesa, e perguntei-lhe se sabia do caso. Que sim; que os soviéticos haviam informado do sucedido e que o corpo da Vera já tinha sido levado *incontinenti* de madrugada para a URSS... e que era melhor eu não fazer perguntas a mais ninguém porque os soviéticos preferiam o silêncio e absoluta discrição... Enfim, coisas deles. A caminho do alojamento na inclemente madrugada islandesa, o táxi telefoniava o Caruso/Donizetti em *Una Furtiva Lagrima*(11)...

E seguem-se umas divagações:

A) No Jornal Oficial L170, de 11.6.2014, p. 1

«*Informação relativa à aplicação provisória de um Protocolo Adicional ao Acordo entre a Comunidade Económica Europeia e o Reino da Noruega na sequência da adesão da República da Croácia à União Europeia*

Tendo ficado concluídos em 15 de maio de 2014 os procedimentos necessários à aplicação provisória de um Protocolo Adicional ao Acordo entre a Comunidade Económica Europeia e o Reino da Noruega (...)»

Muitas vezes a simples utilização do particípio passado — não vale a pena complicar! — permite economizar termos. No exemplo *supra*, poder-se-ia dizer:
«*Concluídos em...*»

B) No Jornal Oficial L170, de 11.6.2014, p. 49

«*Decisão do Conselho, de 19 de maio de 2014, respeitante à celebração do Acordo entre a União Europeia e o Principado do Listenstaine sobre as modalidades da sua participação no Gabinete Europeu de Apoio em matéria de Asilo*
...
(3) Tal como especificado no considerando 21 do Regulamento (UE) n.º 439/2010 do Parlamento Europeu e do Conselho (2), o Reino Unido e a Irlanda participam nesse regulamento e estão a ele vinculados (...)
(4) Tal como especificado no considerando 22 do Regulamento (UE) n.º 439/2010, a Dinamarca não participa nesse regulamento e não está a ele vinculada (...)»

Este «*tal como especificado*» é de uso e abuso inútil, podendo escrever-se, simplesmente, por exemplo:
«*Conforme o considerando...*»
ou
«*Segundo o considerando...*»

C) No Jornal Oficial L 172, de 12.6.2014, p. 28

Decisão do Conselho, de 26 de maio de 2014, relativa à posição a adotar em nome da União Europeia na 103.ª sessão da Conferência Internacional do Trabalho sobre as emendas ao Código da Convenção do Trabalho Marítimo

Na versão EN desta Decisão diz-se, na parte sublinhada, «amendments»; em FR: «amendements»; em ES: «enmiendas»; em DE: «Änderungen»; em IT: «modifiche»; em NL: «wijzigingen».

As «emendas» surgem com frequência em versões portuguesas de documentos internacionais. Há autores que consideram que emendas é o mesmo que alterações.

A Convenção de Viena sobre o Direito dos Tratados de 1969, ratificada por Portugal em 2003[12], não contém o termo «emenda», falando na parte IV, artigo 39.º e segs., em «revisão e modificação», ao passo que a versão francesa fala em «amendement et modification» e a inglesa em «amendment and modification».

A terminologia «emendas» pode ter origem, salvo erro, na Constituição dos EUA de 1787 onde as sucessivas alterações ou aditamentos foram designados por «amendments». Nas versões portuguesas dessa Constituição, os «amendments» foram traduzidos por «emendas»... assim se tendo consagrado tal termo... E isto porque, amiúde, a língua de partida «contamina» a de chegada, traduzindo-se para esta um termo por via de uma simples «transposição» por similitude com o termo usado naquela...

Jorge Miranda, por seu turno, chama «aditamentos»[13] aos «amendments» da Constituição americana, o que parece ser mais correto neste caso.

No caso da Decisão do Conselho acima mencionada, creio que se deveria dizer «as alterações» e não «as emendas», sendo certo, aliás, que não se trata de emendar nada... porque, a não se entender assim, seria pior a emenda que o soneto...

D) Decreto Legislativo Regional n.º 10/2014/A[14]

Artigo 4.º
Liquidação
As pessoas singulares ou coletivas que exerçam as atividades mencionadas na alínea b) do artigo 2.º submetem anualmente aos serviços competentes da Administração Regional uma declaração da qual consta a quantidade de sacos de plástico adquiridos e a quantidade de sacos de plástico distribuídos aos consumidores finais no ano civil anterior, para fins de cálculo da taxa a liquidar, devendo proceder ao seu pagamento num prazo não superior a noventa dias, a contar da data da declaração.

Com frequência os articulados de lei são extensos, a ponto de tirar o fôlego a pessoas de bem, coisa que não se faz e a sociedade desaprova.

No texto acima, embora não seja um daqueles exemplos que assustam pela extensão, introduziria uma paragem a meio (um ponto final) para os leitores poderem ingerir um copo de água ou de outro líquido adequado ou verem um concurso na TV e reduziria o texto, ficando assim (a cinzento a intervenção no texto original):

As pessoas singulares ou coletivas que exerçam as atividades mencionadas na alínea b) do artigo 2.º submetem anualmente aos serviços competentes da Administração Regional uma declaração. Desta consta a quantidade de sacos de plástico adquiridos e ~~a quantidade~~ de sacos de plástico distribuídos aos consumidores finais no ano civil anterior, para fins de liquidação da taxa, devendo aquelas efetuar o pagamento nos noventa dias subsequentes à declaração.

Nota ao que sugiro *supra:*
«*cálculo da taxa a liquidar*» pode reduzir-se a *liquidação da taxa*, dado que nesta liquidação (apuramento do valor a pagar) já se contém o cálculo...Ou estou enganado?

E) Resolução da Assembleia da República n.º 51/2013[15]

No artigo 2.º, Definições, alínea f), pode ler-se, por extrato:

f) «Autoridade competente» significa qualquer serviço de informações, autoridade responsável pela aplicação da lei, serviço de imigração e autoridade de segurança pública de cada uma das Partes, incluindo, para os Estados Unidos (...)

E na correspondente versão EN:
f) «Relevant authority» means any intelligence, law enforcement, immigration and public security authorities of the respective Parties including, for the United States (...).

No sublinhado nas duas versões lá surge em PT *autoridade responsável pela aplicação da lei* como tradução de *law enforcement* (...) *authoritie(s).*
Parece-me uma tradução pouco curial. Esta questão já foi focada n'«a folha» n.º 24[16], pelo que me permito remeter os leitores para o que aí se disse adrede.

Artigo 10.º
Proteção contra a divulgação
(....)
4 — Os pedidos de consentimento para uma divulgação deverão ser feitos segundo os seguintes procedimentos: se a Parte destinatária está interessada em obter autorização para utilizar qualquer informação de rastreio do terrorismo ao abrigo do presente Acordo, em qualquer ação ou processo judicial ou administrativo, qualquer processo judicial ou quase-judicial, ou em qualquer processo que poderia acarretar divulgação pública, deverá primeiro contactar a Parte transmissora através do seu ponto de contacto, o qual envidará esforços no sentido de obter autorização da autoridade competente da qual emana a informação.

Versão em EN:
4 — Requests for consent to a disclosure shall be made under the following procedures: if the Receiving Party is interested in obtaining authorization to use any terrorism screening information provided under this Agreement in any legal or administrative proceeding or process, any judicial or quasi-judicial process, or in any process that could result in public disclosure, the Receiving Party shall first contact the Providing Party through its point of contact, which will endeavor to obtain permission from the relevant authority that originated the information.

Este dispositivo do n.º 4 poderia ter, na versão PT, a seguinte redação alternativa:
4 — Os pedidos de consentimento de divulgação seguem os seguintes procedimentos: se a Parte destinatária pretende autorização para utilizar qualquer informação de rastreio do terrorismo ao abrigo do presente Acordo, em ação ou processo judicial ou administrativo, em processo quase — judicial, ou em processo que poderia acarretar divulgação pública, contacta primeiro a Parte transmissora através do seu ponto de contacto, o qual diligencia por obter autorização da autoridade competente da qual emana a informação.

Uma breve nota: é de suprimir, quando possível, o verbo «dever» nas suas múltiplas declinações a acompanhar um outro verbo (v.g.: *deverá...contactar*). A simples utilização do presente do indicativo no verbo principal arruma a questão da obrigatoriedade (do dever) da ação[17].

F) Venho insistindo na questão das legendas de filmes/séries na TV, particularmente quando são de origem americana e implicam polícias, crimes, tribunais e toda a panóplia jurídica.

Eu sei que, assim como é estultícia ter atrás de cada cidadão português um polícia que impeça que se estacione a viatura sobre o passeio, também é impossível haver um jurista de qualidade atrás do ou no papel de autor das legendas em português, de modo a que não saia uma embrulhada tão imensa que faça chorar o Senhor Jesus da Piedade da simpática cidade raiana de Elvas...

Ainda recentemente vi em legenda na TV uma expressão como «invalidação de promoção», no meio de uma salganhada em tribunal americano.

Creio que, sendo possível, seria mais profícuo pôr isto numa linguagem «jurídica» menos confusa, aproximando-a do «jurídico» português comum, o que, reconheço, não é tarefa de somenos, pois que, no caso apontado, seria aportuguesar linguagem jurídica do sistema americano, profundamente diverso do nosso. Mas aqui deixo novo apelo!

luis.f.sabino@gmail.com

---

(1)  © L.F.P.L.Sabino
População (2014): 325.000; área: 103.000 km²; coordenadas: 65° 0′ 0″ N, 18° 0′ 0″ O.
Ver Diamond, J., *Collapse: How Societies Choose to Fail or Survive. Chapter VI — The Viking prelude and fugues*, Penguin Books, Reino Unido, 2011, ISBN 978-0-241-95868-1.

(2) Recorda o Farrusco, nome de cão. O Farrusco viveu na Bélgica (Aalst 1983 — Bruxelas 1997) e era danado por cadelas. Apontamento familiar que não interessa obviamente a terceiros.

(3) Entre abundante documentação:
— Adelman, K., *Reagan at Reykjavik: Forty-Eight Hours that Ended the Cold War*, HarperCollins Publishers, 2014, ISBN 978-0-06-231019-4;
— Kissinger, H., *Diplomatie*, Fayard, Paris, 1996, ISBN 2-213-59720-0, pp. 712-713;
— Goodby, J. E., «Looking Back: The 1986 Reykjavik Summit», *Arms Control Today*, vol. 36, setembro 2006, Arms Control Association, Washington, https://www.armscontrol.org/act/2006_09/Lookingback.

(4) Para quem se interessa por estas coisas:
— Reagan, R. , «Address to the Nation on the Meetings with Soviet General Secretary Gorbachev in Iceland, October 13, 1986», *Ronald Reagan Presidential Library and Museum*, http://www.reagan.utexas.edu/archives/speeches/1986/101386a.htm;
— «Summit Aftermath: The View from Moscow; Excerpts From Speech by Gorbachev About Iceland Meeting», *New York Times*, 15.10.1986, http://www.nytimes.com/1986/10/15/world/summit-aftermath-view-moscow-excerpts-speech-gorbachev-about-iceland-meeting.html?pagewanted=1.

(5) Wikipédia, *Höfði*, https://pt.wikipedia.org/wiki/H%C3%B6f%C3%B0i.

(6) YouTube, *Les Chansons de Edith Piaf : Non, je ne regrette rien*, HalidonMusic, https://www.youtube.com/watch?v=yGx0RJ3WzrA&list=RDyGx0RJ3WzrA#t=0.

(7) YouTube, *Delfins - Baía de Cascais*, Jorge Soares, https://www.youtube.com/watch?v=SbkQij2hF54:

Na baía de Cascais
Avistei ao longe um barco a arder
Perguntaste porque o sonhava
Olhei ao céu, não pude responder
Eu pinto esta baía assim
E são mil cores ao pé de mim
Nesta baía eu descobri
Tantas imagens perto de mim
Só, no cais
Vou recordar esse teu olhar

à deriva no mar
Lembro o mar nos teus olhos
Ao deixar neste quadro
a saudade, depois de te perder
Eu pinto esta baía assim
E são mil cores ao pé de mim
Nesta baía eu descobri
Tantas imagens perto de mim
Nesta baía eu vi-te assim,
Nesta baía sinto assim.

(8) A propósito da guerra do Vietname pode ler-se também com interesse: McNamara, R. S., VanDeMark, B., *In Retrospect: The Tragedy and Lessons of Vietnam*, Vintage Books, 1996, ISBN 0-679-76749-5.

(9) Ver Acemoglu, D., Robinson, J. A., *Why Nations Fail: The Origins of Power, Prosperity and Poverty*, Crown Business, 2012, ISBN 978-0-307-71921-8.

(10) Carvalho, A. S., «Poema dos Motoristas Oficiais», *Lisboas: Roteiro Sentimental*, Quetzal Editores, Lisboa, 2000, ISBN 972-564-427-1.

(11) YouTube, «Enrico Caruso - Una Furtiva Lagrima (Remastered)», GreekCallas Channel, http://www.youtube.com/watch?v=K4fUAVcXeiQ.

(12) «Decreto do Presidente da República n.º 46/2003 que ratifica a Convenção de Viena sobre o Direito dos Tratados, feita em 23 de maio de 1969, e respetivo anexo», *Diário da República*, n.º 181, I-A Série, de 7 de agosto de 2003, https://dre.pt/pdfgratis/2003/08/181A00.pdf.

[13] Miranda, Jorge, *Manual de Direito Constitucional*, 4.ª edição, Coimbra Editora, 1990, ISBN 72-32-0420-7 — tomo I, p. 138.
[14] «Decreto Legislativo Regional n.º 10/2014/A que cria medidas para a redução do consumo de sacos de plástico e aprova o regime jurídico da taxa ambiental pela utilização de sacos de plástico distribuídos ao consumidor final», *Diário da República*, n.º 126, I.ª série, de 3 de julho de 2014, http://www.dre.pt/pdf1s/2014/07/12600/0368003682.pdf.
[15] «Resolução da Assembleia da República n.º 51/2013que aprova o Acordo entre a República Portuguesa e os Estados Unidos da América para a Troca de Informação de Rastreio do Terrorismo, assinado em Washington em 24 de julho de 2012», *Diário da República*, n.º 71, 1.ª série, de 11 de abril de 2013, p. 2125, http://www.dre.pt/pdf1s/2013/04/07100/0212502133.pdf.
[16] Sabino, L. F. P. L., «A prima Idalina» *in* «a folha», n.º 24 — primavera de 2007, http://ec.europa.eu/translation/portuguese/magazine/documents/folha24_pt.pdf.
[17] Cf. *Regras de Legística a Observar na Elaboração de Atos Normativos da Assembleia da República*, Divisão de Edições da Assembleia da República, Lisboa, outubro de 2008, ISBN 978-972-556-499-8, http://www.parlamento.pt/ArquivoDocumentacao/Documents/AR_Regras_Legistica.pdf.
Ver tb. *Guia Prático Comum do Parlamento Europeu, do Conselho e da Comissão para as pessoas que contribuem para a redação de textos legislativos nas instituições comunitárias*, Serviço das Publicações, Luxemburgo, 2003, http://eur-lex.europa.eu/techleg/index.html?locale=pt.

A Lusofonia — um porta-aviões em mar revolto

Philippe Magnan Gariso
Tradutor técnico — Mota-Engil, Engenharia e Construção, SA, Lisboa

O mar é — sempre foi —, porque não dizê-lo, fonte do vate poético, suporte de metáforas, morada de divindades aquáticas, foco de desentendimentos, cenário de batalhas, conquistas, via para expansionismos e, também, desde tempos imemoriais, motivou compras — algumas legítimas, outras dúbias. Por essa imensidão oceânica, em busca de novas civilizações, culturas, povos, terras, sulcaram desde simples embarcações a gigantescos vasos de guerra. Ora, ao leitor não escapam as imagens recentes de que se socorrem alguns para falar de economia e do potencial da língua portuguesa no mundo; pois é: o porta-aviões; não faltam referências ao porta-aviões para o arranque económico e para tantas outras finalidades. Mas falemos desse potencial, dessa força magnânima que nos une, projecta, cria pontes, que vence o mar… vencerá? A portuguesa língua e a lusofonia.

Atenta a finalidade da lusofonia, é indiscutível que tem de ser forte, sólida, abrangente; é, também, inegável que a espinha dorsal dessa lusofonia, a sua âncora maior, a língua portuguesa, ferramenta veicular por excelência, tem de ser aprumada, vigorosa, assertiva, mas igualmente dúctil e permeável quanto baste para vencer o desafio da inter/intraculturalidade e da pluralidade; despojada a língua-pátria do seu papel de padrão, de modelo, de bússola, a lusofonia não poderá senão esboroar-se. Claro está, contanto que por lusofonia se entenda o espaço multifacetado das artes, de divulgação da ciência e das ideias — espaço despolitizado, mas em contrapartida universalizado — sem a conspurcação do mero exercício do poder ideológico.

Enquanto veículo, a língua portuguesa tem de funcionar como bússola, quer na vertente gramatical quer na semântica ou no vocabulário, isto é, não funcionar como modelo imposto, rígido, estereotipado, mas antes como ponto de referência, como *Norte Magnético* a partir do qual as variantes se desenvolvem, se declinam, mas também a partir do qual essas variantes se balizam — e se sabemos o quão fundamentais são as balizas na contenção dos excessos, desvios, atropelos, incongruências (sobretudo terminológicas e etimológicas).

Ora, o leitor começa a ter a percepção do caminho que seguimos. A língua portuguesa maltratada, mal pronunciada, mal escrita no seu território berço, terras por que se bateram Viriato e Sertório, dificilmente ganhará peso nas instâncias internacionais; nada tenho contra as gentes de Vera Cruz, mas

o leitor tem consciência de que é dessas bandas que nos vem, se não a imposição, pelo menos a tentativa, e gostaria de acompanhar o leitor na crença — ou seja, admitir que o falar das Terras de Vera Cruz ganha terreno nestes lugares aquém- e além-Tejo, por via da pujança do idioma, não por outros poderes. Que fique bem claro, sim, porque se impõe, desde já, a clarificação: não se pode confinar a lusofonia à CPLP (surgida fundamentalmente por impulso de um diplomata brasileiro) sob pena de subjugação da noese e da dianóia (no seu sentido aristotélico) à ideologia.

Não se pode exportar um produto, um bem ou um serviço se descuidarmos o seu valor intrínseco e a respectiva embalagem; quando nas escolas, na comunicação social, nos meios em que se move o escol político, entendendo-se os que administram a *polis*, se desconhecem as regras da pronominalização, da concordância verbal, da conjugação dos verbos, da ortofonia, quando se atropelam as regras básicas da tipografia, em que se inclui a escrita incorrecta das unidades de medida, da abreviação, siglação e acronímia, perde-se a mais-valia que muitos, não sei se por modismo, ou vaidade, apregoam. Uma lusofonia robusta, em que todos caibam, tem de sustentar-se num tronco são, numa guia que norteia os falantes-utilizadores.

Quando a esse tronco, ou casco, se infligem maus tratos, também por via de um acordo que não consubstancia nem resulta de uma **reforma**, desfere-se um golpe funesto nesse poderoso veículo, despojando-o da sua seiva para se lhe sobrepor a vontade de outros que por sinal reconhecem — agora — que terão percorrido este tempo com velocidade excessiva, e outros, noutro continente, aderem ao adiamento. E o leitor pergunta de chofre: «Porque não se começou pelo princípio, ou seja, elaborar um vocabulário técnico comum?». Ora, a inexistência desse vocabulário comum leva a que, em determinadas áreas terminológicas, textos técnicos produzidos em Portugal e no Brasil tenham de ser reciprocamente traduzidos para que os destinatários de ambas as línguas se entendam.

O leitor sabe tão bem quanto eu que, a despeito das diferenças de pronúncia, de ortografia, de organização frásica, nas mais das vezes por razões que se prendem com a gramática, dos cognatos com sentido diferente, os falantes da língua inglesa espalhados pelo mundo (Reino Unido, EUA, África do Sul, Austrália, Nova Zelândia, por exemplo) não precisaram de reformas, nem de acordos para comunicar, nem o poderio de um deles teve de se impor; a prova é que muitos livros de estilo (outrora «manuais de redacção») recomendam a grafia-padrão, ou seja a do inglês britânico.

Em relação à língua francesa, a reforma da ortografia de há quase 20 anos a nada conduziu, ainda que as novas grafias estejam dicionarizadas — nem os manuais escolares, nem os professores acataram peremptoriamente.

Quanto ao alemão, a reforma ortográfica que foi tornada obrigatória a partir de 2005 e reuniu à mesa especialistas alemães, austríacos, helvéticos e luxemburgueses (estes últimos não aderindo), é torpedeada por importantes grupos de imprensa escrita e alvo de críticas.

Finalmente, a hispanofonia não precisou dessa genial reforma ou acordo para se manter una e dinâmica e para fomentar a aprendizagem da língua, embora algumas alterações tenham sido introduzidas.

Pois é, ao leitor se depara desde já uma conclusão: rescrever, suprimir ou acrescentar consoantes, e suprimir acentos não se impõe — as palavras carregam história, contam-nos história e não se desfaz a história com um despacho ou em gabinetes restritos. Não é a partir de um gabinete que se faz explodir o porta-aviões, sobretudo o da lusofonia.

Quando se fala em divulgar o ensino da língua portuguesa no estrangeiro, e estou a pensar na Rússia e na China, por ter sido alvo de uma recente reportagem na imprensa falada, na primeira linha de professores estão os falantes de terras de Vera Cruz, fenómeno, aliás, antigo, já que em França, por exemplo, nos anos 80, o leitorado português era brasileiro! E o leitor tem consciência de que a pronúncia não é a pronúncia-padrão, o vocabulário é diferente, o modelo gramatical em muitos casos diverge, logo os aprendentes aprendem algo muito diverso do padrão. Não que isso conflitue com a

diversidade, a criatividade, a modernidade, mas a base, o tronco, o casco, deslocou-se não de um país para outro, mas de uma cultura para outra, de um padrão para uma variante e é este pecado que desvirtua, pauperiza a mais-valia, o tal porta-aviões, cujo casco, ao sabor do mar revolto, se vai fragilizando; aliás, o leitorado brasileiro tem como finalidade expandir **a língua e a cultura brasileiras**! Alguns sítios electrónicos portugueses sobre a língua pátria falam do português como língua de ciência, mas não fazem mais do que evocar alguns escritores de há dois ou três séculos; então e a modernidade, o progresso, as instâncias internacionais? Na FAO, no FNUAP, na OCDE, na OMS, na OIT, na OTAN, na UNESCO, no Secretariado-Geral da ONU ou na UIP, por exemplo, não encontramos o português como língua de trabalho, nem como língua oficial; imagine-se, nem na FIFA. A excepção é a SADC (Comunidade para o Desenvolvimento da África Austral).

Quantas vezes, senão diariamente, o tal porta-aviões é bombardeado, torpedeado por uns quantos — muitos — submarinos com estrangeirismos, sobretudo na imprensa falada, sinal, pensará, e muito acertadamente, o leitor, da preguiça intelectual, do desinteresse em encontrar termos e expressões equivalentes, em suma, da inércia e do desapego do escol em preservar e mostrar aos «stakeholders» e «players» (actores na cena internacional) que a língua portuguesa é dinâmica e vigorosa, e suficientemente elástica para se adaptar a novos conceitos.

Tal como eu, o leitor tem a tentação de estabelecer comparações e, de facto, não posso deixar de pensar na *Direction générale à la langue française et aux langues de France* e nas comissões de terminologia e neologia que existem nos ministérios, em França, cuja missão é enriquecer o vocabulário francês nas várias áreas do saber, fomentar o uso da língua e desenvolver a francofonia; as listas de termos adoptados são publicadas no equivalente francês do *Diário da República*. A esta missão associam-se os parceiros da francofonia, nomeadamente o *Office québécois de la langue française*, o *Bureau de traduction des services gouvernementaux* do Canadá, a *Section de terminologie de la Chancellerie fédérale* na Suíça, o *Service de la langue française pour la Communauté française de Belgique* e a não menos importante *Agence universitaire de la Francophonie* com sede no Canadá.

Caro leitor, não se pretende nem polémica, nem celeuma, apenas um olhar quiçá diferente. Não se pretende afundar nem o porta-aviões nem inutilizar os submarinos com cargas de profundidade; tão só, que as águas sejam «flat» (uma homenagem aos surfistas); que se apaziguem para que todos possamos navegar — uns pachorrentamente, outros com mais fervor e entusiasmo, mas com um único destino: a Lusofonia.

Estou certo de que o leitor se solidarizará numa palavra de grande apreço pela gigantesca tarefa de todos quantos defendem o português na IATE (base terminológica da UE) e de todos quantos nas instituições europeias lidam com a arte e o ofício da tradução em relação aos textos que hajam de ser vertidos de e para português.

philippe.gariso@gmail.com

Bulgária — ficha de país

Luís Seabra
Paulo Correia
Direção-Geral da Tradução — Comissão Europeia

Nesta ficha de país reúne-se informação terminológica relativa à Bulgária que se encontra dispersa por vários documentos normativos ou de referência.

A ficha de país da Bulgária tem um desafio adicional — a língua búlgara utiliza um **alfabeto cirílico**. Para os nomes próprios não traduzidos recorrer-se-á às adaptações mínimas necessárias da transliteração orientada para a língua inglesa de forma a garantir que o leitor lusófono reproduzirá o nome com uma pronúncia o mais próxima possível da língua de origem. Para tanto, também se introduzirá acentuação sempre que não for evidente a localização da sílaba tónica.

Apresenta-se em anexo a esta ficha uma tabela com o alfabeto búlgaro e os respetivos equivalentes usuais em inglês e aproximados em português. A tabela inclui igualmente transcrições fonéticas e a transliteração ISO 9:1995.

REPÚBLICA DA BULGÁRIA (IATE: 860994)

CAPITAL:	Sófia
GENTÍLICO/ADJETIVO:	búlgaro/a (plural: búlgaro/a(s))
MOEDA:	lev (plural: leves)
SUBDIVISÃO:	stotinka (plural: stotinkas)

Principais cidades:	Sófia, Plovdiv, Varna, Burgas, Russe(1)
Rios:	Danúbio, Struma, Maritsa, Iskar
Lagos:	Burgas
Serras:	Stara Planiná(2), Pirin, Mussala

Subdivisões administrativas

#	búlgaro	português	inglês	IATE
2	район	região	region	—
6	район за планиране	região do plano	planning region	—
28	област	distrito	province	3502196
264	община	município	municipality	3553011
5329	населено място	aldeia	settlement	—

Fonte: Eurostat, *Nomenclature of territorial units for statistics: National Structures (EU)*, http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/correspondence_tables/national_structures_eu

Regiões

NUTS	БЪЛГАРИЯ	BULGÁRIA	BULGARIA	IATE
BG3	**Северна и Югоизточна България**	**Bulgária do Norte e Sudeste**	**North and South-East Bulgaria**	—
BG31	Северозападен	Noroeste	North-West	—
BG311	Видин	Vidin	Vidin	—
BG312	Монтана	Montana	Montana	—
BG313	Враца	Vratsa	Vratsa	—
BG314	Плевен	Pleven	Pleven	—
BG315	Ловеч	Lóvetch	Lovech	—
BG32	Северен централен	Norte Central	North-Central	—
BG321	Велико Търново	Grande Târnovo	Veliko Tarnovo	—

BG322	Габрово	Gábrovo	Gabrovo	—
BG323	Русе	Russe	Ruse	—
BG324	Разград	Razgrad	Razgrad	—
BG325	Силистра	Silistra	Silistra	—
BG33	Североизточен	Nordeste	North-East	—
BG331	Варна	Varna	Varna	—
BG332	Добрич	Dóbritch	Dobrich	—
BG333	Шумен	Chúmen	Shumen	—
BG334	Търговище	Targóvichte	Targovishte	—
BG34	Югоизточен	Sudeste	South-East	—
BG341	Бургас	Burgas	Burgas	—
BG342	Сливен	Sliven	Sliven	—
BG343	Ямбол	Iambol	Yambol	—
BG344	Стара Загора	Velha Zagora	Stara Zagora	—
BG4	**Югозападна и Южна Централна България**	**Bulgária do Sudoeste e Sul Central**	**South-West and South-Central Bulgaria**	—
BG41	Югозападен	Sudoeste	South-West	—
BG411	София (столица)	Sófia (capital)	Sofia (capital)	—
BG412	София	Sófia	Sofia	—
BG413	Благоевград	Blagoevgrad	Blagoevgrad	—
BG414	Перник	Pernik	Pernik	—
BG415	Кюстендил	Kiustendil	Kyustendil	—
BG42	Южен централен	Sul Central	South-Central	—
BG421	Пловдив	Plovdiv	Plovdiv	—
BG422	Хасково	Háskovo	Haskovo	—
BG423	Пазарджик	Pazardjik	Pazardzhik	—
BG424	Смолян	Smolian	Smolyan	—
BG425	Кърджали	Kârdjali	Kardzhali	—

Fonte: Serviço das Publicações, *Código de Redação Interinstitucional: Anexo 10 — Lista das Regiões*, http://publications.europa.eu/code/pt/pt-5001000.htm;
Correia, P., «Regiões dos 27 — Bulgária e Roménia» *in* «a folha» n.º 26 — primavera de 2008, http://ec.europa.eu/translation/bulletins/folha/folha26.pdf

Órgãos judiciais

#	**búlgaro**	**português**	**inglês**	IATE
113	районен съд	tribunal de comarca(3)	district court	2244257
28	окръжен съд	tribunal distrital(4)	provincial court	2244301
28	административен съд	tribunal administrativo	administrative court	—
5	апелативен съд	tribunal de recurso	court of appeal	—
5	военен съд	tribunal militar	military court	—
1	военно-апелативен съд	tribunal militar de recurso	military court of appeal	3544429
1	върховен административен съд	supremo tribunal administrativo	supreme administrative court	3544414
1	върховен касационен съд	supremo tribunal de cassação	supreme court of cassation	3544413
1	конституционен съд	tribunal constitucional	constitutional court	3544412

Fonte: Portal Europeu da Justiça, *Sistemas judiciais nos Estados-Membros — Bulgária*, https://e-justice.europa.eu/content_judicial_systems_in_member_states-16-bg-pt.do?member=1.

Luis.Seabra@ec.europa.eu
Paulo.Correia@ec.europa.eu

Anexo: Alfabeto búlgaro

Esta tabela é idêntica à que já foi publicada como anexo ao artigo «Transliteração do cirílico (búlgaro)... a partir do inglês»(5).

letra búlgara	fonética	ISO 9[6]	romanização inglesa[7]	equivalente português[8]
Аа	/a/ ou /ɐ/	a	a	a
Бб	/b/	b	b	b
Вв	/v/	v	v	v
Гг	/g/	g	g	g, gu (antes de e/i)
Дд	/d/	d	d	d
Ее	/ɛ/	e	e	e
Жж	/ʒ/	ž	zh	j
Зз	/z/	z	z	z
Ии	/i/	i	i	i
Йй	/j/	j	y	i (breve)
Кк	/k/	k	k	k
Лл	/l/ ou /ɫ/	l	l	l
Мм	/m/	m	m	m
Нн	/n/	n	n	n
Оо	/ɔ/ ou /o/	o	o	o
Пп	/p/	p	p	p
Рр	/r/	r	r	r
Сс	/s/	s	s	s, ss (intervocálico)
Тт	/t/	t	t	t
Уу	/u/ ou /o/	u	u	u
Фф	/f/	f	f	f
Хх	/x/	h	h	h[9]
Цц	/t͡s/	c	ts	ts
Чч	/t͡ʃ/	č	ch	tch
Шш	/ʃ/	š	sh	ch
Щщ	/ʃt/	št	sht	cht
Ъъ	/ɤ/ ou /ɐ/	ă	a	a
Ьь	/j/ ou não pronunciado	'	y	i
Юю	/ju/ ou /jo/	û	yu	iu
Яя	/ja/ ou /jɐ/	â	ya	ia

Em resumo:

IA e IU para transliterar Я e Ю, respetivamente (não YA ou YU)
I para transliterar Й ou Ь (não Y)
CH para transliterar Ш (não SH)
CHT para transliterar Щ (não SHT)
TCH para transliterar Ч (não CH)
J para transliterar Ж (não ZH)
GU antes do E e I (não G)
SS dobrado para C quando entre vogais (não S)

---

(1) O nome de família da atual presidente do Brasil, **Dilma Rousseff**, está ligado ao nome da cidade de Русе. Русев — **de Russe**. Se Петър Русев (naturalizado Pedro Rousseff), o pai de Dilma Rousseff, tivesse emigrado hoje para o Brasil, o nome romanizado que constaria no passaporte búlgaro seria **Petar Rusev**. Em finais dos anos trinta do século XX utilizava-se a transliteração francesa: **Rousseff**. Apenas nos anos noventa os documentos de viagem búlgaros passaram a utilizar uma transliteração orientada para o inglês.
(2) Literalmente: Velha Montanha.
(3) Район indica tanto uma divisão estatística como uma divisão judicial, não coincidindo geograficamente (há duas regiões estatísticas NUTS 1 e 113 comarcas e respetivos **tribunais de comarca**).
(4) Окръг é um distrito judicial, област é um distrito administrativo. Não coincidem geograficamente.
(5) Correia, P., «Transliteração do cirílico (búlgaro)... a partir do inglês», *in* «a folha», n.º 31 — outono de 2009, http://ec.europa.eu/translation/portuguese/magazine/documents/folha31_pt.pdf.
(6) ISO 9:1995 — *Transliteration of Cyrillic characters into Latin characters. Slavic and non Slavic languages*. Existe, ainda, a Norma Portuguesa NP-47 (1961) — *Sistema internacional para a transliteração dos carateres cirílicos* (desatualizada).
(7) Regulamento n.º 13, de 8 de fevereiro de 2012, de emissão de documentos pessoais búlgaros. Tabela para a transliteração do alfabeto búlgaro para o alfabeto inglês, http://lex.bg/bg/laws/ldoc/2135663268.

Existem inúmeros outros sistemas de transliteração (romanização) do alfabeto búlgaro orientados para o inglês. Exemplos: ALA/LC (American Library Association/Library of Congress); BGN/PCGN (United States Board on Geographic Names/Permanent Committee on Geographical Names for British Official Use); Dantchev; etc.
Consultar igualmente o anexo 2 (Transliteration table for Cyrillic) do *English Style Guide* da Direção-Geral da Tradução da Comissão Europeia, http://ec.europa.eu/translation/writing/style_guides/english/style_guide_en.pdf.

(8) É conveniente notar que se trata aqui de transliterações (manutenção de endónimos) e não de aportuguesamentos (criação de exónimos). Nos aportuguesamentos há algumas regras ortográficas a observar, mas que por vezes são esquecidas:

> «Em vocábulos originários de línguas com alfabetos exóticos (entenda-se: todos os alfabetos que não forem o alfabeto latino — o grego, o cirílico, o hebraico, o japonês, etc.) NUNCA usaremos o CH, os SS e o G, mas sim o X, o Ç e o J.»
> Moreno, C., «Quando se usa o J, o X e o Ç?», *Sua Língua*, http://wp.clicrbs.com.br/sualingua/2009/05/18/quando-se-usa-o-j-o-x-e-o-c/.

(9) «H» aspirado, pronunciado aproximadamente como o j espanhol.

O tsar traído

Jorge Madeira Mendes
Direção-Geral da Tradução — Comissão Europeia

Costumo dizer que falo *x* línguas e meia.
O *x* é o número daquelas que, de um modo geral, todos (mais ou menos) dominamos nos serviços de tradução das instituições europeias — francês, inglês, espanhol... — além, obviamente, da língua-mãe.

E a «meia língua» que falo? Bom, quantificá-la em 50% é talvez exagero — que o meu russo (pobrezito) andará se calhar abaixo (bem abaixo) dos 10%.

Mas o importante é que os conhecimentos incipientes que tenho da língua russa (obtidos, no essencial, por autodidatismo... e alguma intimidade com elementos da comunidade russófona de Bruxelas) são, ainda assim, suficientes para me permitir ler, com pouca ou nenhuma ambiguidade, quaisquer palavras grafadas em cirílico. Isto porque o russo (a par de outras línguas eslavas, inclusivamente as que se representam por carateres latinos) é fonético: quem conhece o alfabeto e as regras ortográficas sabe pronunciar as palavras com bastante exatidão... mesmo que ignore completamente o que está a ler (e, vamos lá, sem fazer grandes exigências quanto à acentuação).

A título de exemplo: quem circular por estradas da Rússia e se deparar com o dístico СТОП, saberá (se conhecer o alfabeto cirílico) que está diante de uma obrigação de STOP. E se, num jornal russo, vir as palavras **фадо** e **Амалия**, saberá que o artigo se refere a *fado* e *Amália*.

A «culpa» destas elucubrações é o artigo do Paulo Correia «Сочи — um estudo de caso da transliteração do russo»[1].
Artigo pertinente — porque alerta para a ausência total de critérios na transliteração dos carateres cirílicos para português.

Consideremos, por exemplo, a cidade de **Хабаровск**, no Extremo Oriente russo: se não fizer a menor ideia do que está a escrever (o que, pelo que se constata na literatura, ocorre na maioria dos casos), cada jornalista ou tradutor português grafará este nome conforme a língua de grafia latina pela qual ele lhe chega: *Khabarovsk* se a fonte for anglo-saxónica, *Jabárovsk* se espanhola, *Chabarowsk* se alemã.
A atitude mais sábia seria a rápida pesquisa de uma pista: se o texto-fonte for espanhol, é de esperar que qualquer jornalista ou tradutor português saiba que a letra «J» não tem, nesta língua, o mesmo valor fonético que em português; se a fonte for o alemão, saberá (ou terá a obrigação de saber) que a

letra «w» representa a consoante que nós grafamos pelo carater «v»; e suspeitará também que, em alemão, o par de letras «Ch» há de representar algo distinto do nosso «Ch» de «Chamusca».
Com as luzes (poucas que sejam) que qualquer jornalista ou tradutor português deve ter destas línguas, e com um expedito cruzamento de informações quanto às respetivas grafias, é provável que chegue à conclusão de que a consoante inicial (representada pelo carater cirílico **X**) é aquele fonema gutural, inexistente em português, que em espanhol se representa pela letra «j» (como em *Javier*, *bajito*, *ajos*, *junio*) e em alemão pelo par «ch» (como em *acht*, *noch*, *auch*, *Frucht*), fonema esse que, também inexistente em inglês, é normalmente representado nesta língua pelo par «kh» (talvez porque foneticamente se situa entre um «k» e um «h» aspirado). Mas, claro, tal investigação exigiria trabalho, pelo que se opta pela lei do menor esforço... e se reproduz a grafia utilizada no texto-fonte, com todos os erros que isso pode ocasionar.

A ausência de uma regra de transliteração dos carateres cirílicos para português está na base de muitos equívocos, como, por exemplo, *czar*: esta palavra deriva de uma transliteração húngara do russo **царь** ou do búlgaro **цар**, que se pronunciam [tsar](2). Acriticamente, adotou-se aquela grafia em português (pelo menos no de Portugal), em detrimento de uma pronúncia mais fiel ao étimo russo. Ora, foneticamente, a palavra «portuguesa» *czar* nada tem a ver com os étimos russo ou búlgaro nem tampouco com as versões nas línguas leste-europeias de grafia latina para as quais aqueles étimos começaram por ser transliterados. Tem-no, sim, a forma (infelizmente rara) *tsar*. Note-se que o francês, o inglês e o português do Brasil (realce propositado) utilizam *tsar* (ainda que qualquer deles registe ocorrências de *czar*). Segundo o *Dicionário Houaiss da Língua Portuguesa*, a forma preferencial é «tsar», embora seja comum encontrar as grafias «czar» ou «tzar».

Importaria adotarmos uma grafia que, segundo as regras do nosso idioma, permitisse ao leitor português médio reproduzir o mais fielmente possível a fonética das línguas em questão.

E regresso assim ao artigo do Paulo Correia.
Deliberadamente, não propõe grafias. O seu objetivo é, tão-só, alertar para as incongruências da imprensa escrita portuguesa quando se trata de transliterar palavras originalmente grafadas em cirílico.

Atrevo-me, pois, a propor uma tabela de transliteração (em anexo), começando por ilustrá-la com os nomes que nesse artigo se citam (e mais alguns).

1) O líder da Revolução de 1917:

- em russo: **Владимир Ильич Ульянов (Ленин)**;
- em português: «Vladímir Ilitch Ulyánov (Lénin)».

Notas:
i) para o seu verdadeiro apelido, seria igualmente aceitável a grafia «Uliánov»;
ii) para o nome que o celebrizou, valeria também a grafia «Lénine», ou até «Lênin» ou «Lênine», estas duas últimas reproduzindo ainda mais fielmente a fonética russa.

2) A cidade que albergou os Jogos Olímpicos de Inverno de 2014:

- em russo: **Сочи**;
- em português: «Sótchi», ou até «Sôtchi», que reproduziria ainda mais fielmente a fonética russa (mas jamais «Sóchi» ou, muito menos, «Sochi», que dela se afastam).

3) O último Chefe de Estado da União Soviética:

- em russo: **Михайл Сергеевич Горбачёв**;
- em português: «Mikhaíl Serghêyevitch Gorbatchov».

Impõem-se algumas explicações:
i) no nome próprio, o par de letras «kh» (correspondente ao carater cirílico **х**) representa a consoante gutural que em espanhol se grafa com a letra «j» e em alemão com o par «ch»: dada a inexistência de tal consoante em português, seguimos a prática inglesa: representação pelo par de letras «kh»;
ii) ainda no nome próprio, o acento no segundo «i» é importante, visto ser sobre essa vogal que recai a tónica;
iii) no patronímico, o par de letras «gh» que precede a vogal «e» representa a consoante gutural que, por exemplo, ocorre duas vezes na palavra portuguesa *gago*; se, com efeito, pretendemos tornar a leitura inequívoca para um português médio, não poderíamos deixar que à letra «g» se sucedesse simplesmente a vogal «e», dado o risco de confusão com a fricativa de *geral*; mas também me parece inconveniente associar-lhe a auxiliar «u» (como em *guerra*), dado o risco de se pensar que se leria «gu»;
iv) o apelido não deve grafar-se «Gorbachev», como em inglês, pois, em primeiro lugar, não é evidente para um leitor português que o par «ch» deve ler-se «tch»;
v) a outra razão por que o apelido não deve grafar-se «Gorbachev» refere-se à terminação «ev», que é errónea, mesmo segundo as regras do inglês, porquanto o carater cirílico **ё** representa o ditongo progressivo «yô»; acontece é que, na escrita do russo, se omite muitas vezes o trema do carater **ё**, confundindo-o com **е**, um outro carater cirílico que representa o ditongo progressivo «yé» (ou, melhor, «yê»); por isso, se transliterarmos através da letra latina «e», estaremos a falsear a pronúncia;
vi) pode levantar-se uma outra questão: se o carater cirílico **ё** representa o ditongo progressivo «yô», por que não escrever «Gorbatchyov» ou «Gorbatchiov», em vez de «Gorbatchov»: a razão é que, segundo a lógica do russo, a semivogal do ditongo «yô» é absorvida (neutralizada) pela consoante fricativa «tch» que a precede.

4) A cidade ucraniana que sofreu um grave acidente numa central nuclear em 1986:

- <u>em russo</u>: **Чернобыль**;
- <u>em português</u>: «Tchernóbil» ou, melhor ainda, «Tchernôbil», mas nunca «Chernobil», que não passa de uma adaptação acrítica da grafia inglesa (substituiu-se apenas o «y» inglês pelo mais português «i», mas, por outro lado, ignorou-se que o par «Ch», embora translitere corretamente a fonética original para inglês, não o faz para português).

Duas outras notas se impõem:
i) representamos por «i» a vogal cirílica **ы**, porque, na verdade, se trata de um fonema algo estranho em português (entre «i», «ê», «a» de *cada*, «e» de *me*, *te*, *se*...); «i» não é a representação perfeita, mas sim uma das melhores, dentro do princípio de levar o leitor português médio a reproduzir o mais fielmente possível a fonética original;
ii) tal como se disse na nota de rodapé 2, o carater **ь** (designado «myágkiy znak» e que é mais propriamente um diacrítico do que uma letra) palataliza ligeiramente a consoante cirílica **л**, equivalente da latina «l» (aproximando-a, por exemplo, do «lh» que ocorre na palavra portuguesa *lhe*); sem o carater **ь**, a consoante **л** tornar-se-ia mais cheia (como o «l» catalão ou americano); por motivos de simplificação, podemos, porém, ignorar essa diferença.

5) O grande poeta russo da era romântica:

- <u>em russo</u>: **Пушкин**;
- <u>em português</u>: «Púshkin» ou «Púshkine».

A letra cirílica **ш** representa a consoante fricativa seca que ocorre nas palavras portuguesas *taxa* e *acho*, mas parecem-me de evitar as grafias «Púxkin» (que poderia levar o leitor português médio a pronunciar «Púkskin») ou «Púchkin» (que poderia levar o mesmo leitor a pronunciar «Pútchkin», por julgar tratar-se de uma transliteração inglesa); portanto, considero que o par de letras «sh», embora não típico em português, é preferível para transliterar a fricativa aqui em questão: o leitor português médio conhece, sem dúvida, o seu valor fonético (é-nos familiar a palavra *Shell*, que em geral pronunciamos corretamente). O objetivo é evitar ambiguidades.

6) O célebre romance de Boríss Pasternak:

- em russo: **[Доктор] Живаго**;
- em português: «[Doutor] Jivago».

Impõe-se uma nota sobre a letra cirílica **ж** (maiúscula: **Ж**), que representa a consoante fricativa doce existente nas palavras portuguesas *já*, *gere*, *agiu*, *Jonas*, *júri*: a melhor forma de a transliterarmos para português é através da letra «j», visto que, no nosso idioma, esta letra representa sempre a dita fricativa doce, independentemente da vogal que se lhe seguir, ao passo que a letra «g» correria o risco de assumir o valor gutural de *gago* ou *gume*. No entanto, tal como em relação à letra cirílica **ш** optei pela representação «sh», em lugar da ambígua «ch» (atendendo a que o leitor português poderia julgar estar em presença de uma transliteração inglesa e pronunciá-la, erroneamente, «tch»), levantar-se-á uma dúvida idêntica neste caso: não poderá o leitor português julgar estar em presença de uma transliteração inglesa e, erroneamente, pronunciar o «j» como em *John*, *jingle*, etc.?(3) A questão é que, neste caso, não temos alternativa: para evitar a confusão com «tch», podemos usar a transliteração inglesa «sh», um par de letras que, embora inexistente em português, é reconhecido pelo leitor português médio como tendo o valor fonético do «ch» de *acho* ou do «x» de *taxa*; mas, para evitar a confusão com «dj», não podemos recorrer à transliteração habitual do carater cirílico **ж** para inglês, que adota o par de letras «zh» (com efeito, o título do romance em inglês é «Doctor Zhivago»), porque o par de letras «zh», mais do que inexistente em português, tenderia a ser associado ao fonema fricativo linguodental de *zero* ou *azar*. Portanto, no caso do carater **ж**, a única forma de o transcrevermos para português é pela letra «j», esperando que o leitor lusófono médio não perca a noção de que deve lê-la segundo as regras de nenhuma outra língua senão a sua.

7) O dirigente soviético que sucedeu a Stálin:

- em russo: **Хрущёв**;
- em português: «Khrushtchov».

Notas:
i) a letra cirílica **щ** representa um fonema inexistente noutras línguas: algo como um «sh» duplo ou arrastado; parece-me aceitável adotarmos para português a transliteração tradicional inglesa: «shtch» (em que o fonema «t» é muito leve);
ii) tal como em «Gorbatchov», não se afiguram aceitáveis as transliterações com a terminação em «ev», porque a letra cirílica **ë** representa o ditongo progressivo «yô». Poderá, também aqui, perguntar-se por que não se translitera este nome como «Khrushtchyov»: uma vez mais, a razão é que, segundo a lógica do russo, a semivogal inicial do ditongo «yô» é absorvida (neutralizada) pela consoante fricativa «shtch» que a precede;
iii) ocorre também que o carater **ë** seja transliterado como «yó» ou «ió»: no primeiro caso («yó»), trata-se apenas de uma ligeira variação de sensibilidade fonética em relação a «yô» (onde uns ouvem «ô», outros ouvirão «ó»); já a forma «ió» me parece de evitar, porque, como se trata de um ditongo progressivo, o fonema inicial é uma semivogal, que é mais correto representar pela letra latina «y».

8) O presidente da Ucrânia recentemente deposto:

- em russo: **Янукович**;
- em português: «Yanukóvitch» ou «Yanukôvitch».

Nota: a letra cirílica **я** (maiúscula: **Я**) representa o ditongo «yá»; também aqui me parece de evitar a forma «iá», porque, como se trata de um ditongo progressivo, o fonema inicial é uma semivogal, que é mais correto representar pela letra latina «y».

Jorge-Madeira.Mendes@ec.europa.eu

Tabela de transliteração do alfabeto russo (proposta)

MAIÚSCULAS	minúsculas	transliteração
А	а	a
Б	б	b
В	в	v
Г	г	g (gutural), gh (antes de «e» ou «i»)
Д	д	d
Е	е	ye, yê, yé
Ж	ж	j (como em *já*)
З	з	z
И	и	i
Й	й	i (semivogal, como em *vai*)
К	к	k
Л	л	l
М	м	m
Н	н	n
О	о	o
П	п	p
Р	р	r
С	с	s, ss (posição intervocálica)
Т	т	t
У	у	u
Ф	ф	f
Х	х	kh (gutural seca arrastada)
Ц	ц	ts
Ч	ч	tch
Ш	ш	sh
Щ	щ	shtch (o «t» é quase inaudível)
Ы	ы	i (breve, como «a» de *cada*)
Ь	ь	’ (sinal de palatização)
Э	э	e (aberto, como «é»)
Ю	ю	yu, yú, iú, iu
Я	я	ya, yá, iá
Ё	ё	yo, yô, yó, io, iô, ió

---

[1] Correia, P., «Сочи — um estudo de caso da transliteração do russo » *in* «a folha» n.º 44 — primavera de 2014, http://ec.europa.eu/translation/portuguese/magazine/documents/folha44_pt.pdf.

[2] Há uma pequena diferença de pronúncia entre a palavra russa e a búlgara: na primeira, o carater **ь** (designado «myágkiy znak» e que é mais propriamente um diacrítico do que uma letra) palataliza ligeiramente a consoante cirílica **р** (equivalente da latina «r»). Essa diferença é, porém, desprezável para ouvidos portugueses.

[3] Pior, se bem que mais improvável, seria a possibilidade de o leitor português julgar estar em presença de uma transliteração espanhola e, erroneamente, pronunciar o «j» como em *Javier*, *bajito*, *ajos*, *junio*.

Do milhardo perdido ao caso dos mil milhões

Paulo Correia
Comissão Europeia — Direção-Geral da Tradução

> **bilhão** *m.* O mesmo ou melhor que **billião**. Mil milhões, (1.000.000.000), segundo o systema francês; *ou* um milhão de milhões *ou* um conto de contos, (1.000.000.000.000), segundo o systema inglês. (Fr. *billion*. Cp. *milhão*)
>
> Figueiredo, C. de, *Novo Diccionário da Língua Portuguesa*, 1913(1)

No n.º 18 d'«a folha» em «Em torno do bilião»(2) já houve a oportunidade de falar da incerteza associada à nomenclatura dos grandes números, dada a utilização nos países ocidentais de **duas regras** concorrentes que recorrem exatamente aos mesmos termos para designar quantidades muito diferentes. Assim, por exemplo, conforme a regra utilizada, um **bilião** tanto se pode referir a mil milhões como a um milhão de milhões. Isto é, tanto se pode passar para uma nova designação quando se acrescentam mais três algarismos a um número — **escala curta** — como quando se acrescentam mais seis algarismos — **escala longa**.

	1.000.000 (10^6)	**1.000.000.000** (10^9)	**1.000.000.000.000** (10^{12})
escala curta (regra 3n)(3)	*milhão*	*bilião*	*trilião*
escala longa (regra 6N)(4)	milhão	mil milhões	bilião

Em **Portugal**, antes do século XIX, utilizava-se tradicionalmente a escala longa (regra 6N) — o conto(5). Na última metade do século XIX e primeira metade do século XX, em Portugal conviviam o antigo sistema francês (regra 3n) e o antigo sistema inglês (regra 6N), com claro predomínio do sistema francês(6). Esse predomínio do antigo sistema francês em Portugal (e no Brasil) terá ficado a dever-se à **tradução** para português de obras científicas francesas por ocasião da introdução do sistema métrico decimal em Portugal, a partir de **1852**.

A Europa continental encontrava-se então dividida entre a escala curta (França, Itália, Bélgica) e a escala longa (Reino Unido, Alemanha, Espanha, Países Baixos). Porém, com a vulgarização dos grandes números, devida ao aumento dos valores da produção e do consumo a nível mundial e, sobretudo, ao fenómeno da **hiperinflação**(7) no período entre as duas guerras mundiais, foi sentida a necessidade de uma harmonização na denominação dos grandes números, pelo menos ao nível da Europa. Em **1948**, na IX Conferência Geral dos Pesos e Medidas(8), foi recomendado que os países europeus adotassem a escala longa, alinhando-se, assim, pelo sistema tradicional do Reino Unido.

Após 1948, os países europeus foram adotando a recomendação da IX Conferência Geral dos Pesos e Medidas. A França passou oficialmente à escala longa em 1961(9), tendo-se generalizado o uso de *milliard* para designar mil milhões. No entanto, em sentido contrário, o Reino Unido abandonou a escala longa, alinhando-se progressivamente com os Estados Unidos(10), que se haviam mantido fiéis ao antigo sistema francês(11) (*billion* = mil milhões) adotado após se terem tornado independentes do Reino Unido (também mudaram o sentido de circulação rodoviária para a direita).

países-chave	**pré-1948**	**pós-1948**
França	regra 3n	regra 6N (alinhando c/ Reino Unido…)
Reino Unido	regra 6N	regra 3n (…alinhando c/ Estados Unidos)
Estados Unidos	regra 3n (alinhando c/ França)	regra 3n

Em **Portugal**, a escala longa (regra 6N) foi oficialmente (re)adotada por meio das portarias n.os 14 608 e 17 052(12), respetivamente de 11 de novembro de 1953 e 4 de março de 1959, e fixada na norma portuguesa NP 18:1960 (Ed. 2)(13), reconfirmada em 2006 na NP 18:2006 (Ed. 3)(14). No Brasil continuou a utilizar-se a escala curta.

A prática na tradução e na imprensa

Quando nas **instituições europeias**, nos anos oitenta e inícios dos anos noventa, se traduzia maioritariamente do francês (já então na escala longa), vários tradutores defendiam a criação do neologismo «milhardo»[15] como tradução do francês ***milliard***, para designar 1 000 000 000 (isto é, 10^9). A proposta não vingou, continuando-se a utilizar «mil milhões», como está aliás estabelecido no *Código de Redação Interinstitucional*[16].

No final dos anos noventa e sobretudo no início do século XXI os textos a traduzir nas instituições europeias passaram a ser quase exclusivamente em inglês. Desde então, a questão do «milhardo» nunca mais foi levantada. O problema passou a ser a tradução do ***billion*** «americano» (10^9), pois também nos textos em língua inglesa das instituições europeias se utiliza agora a escala curta.

> ***Billion***. *The use of billion to designate thousand million (rather than million million) is now officially recognised by the Commission and is standard usage in official EU publications. Leading British newspapers and journals (such as the Financial Times and The Economist) have also adopted the convention.*
>
> Comissão Europeia: Direção-Geral da Tradução, *English Style Guide*[17]

Novamente, face ao *billion*, os tradutores da Direção-Geral da Tradução (DGT) mantiveram os «mil milhões», resistindo ao decalque «bilião». Também a generalidade da comunicação social portuguesa utiliza a escala longa, pelo menos quando os textos não resultam da tradução de despachos.

> O corte de **4 mil milhões de euros** na despesa que o Governo irá apresentar à *troika* pretende «eliminar redundâncias» nas funções do Estado, mas não chega para que o esforço do lado da despesa atinja os dois terços.[18]

Em Portugal, o decalque da escala curta — a escala americana — parece restringir-se sobretudo a economistas e alguns políticos[19]. No entanto, o decalque sempre vai aparecendo na imprensa:

> Embora vários economistas critiquem uma austeridade excessiva, o primeiro-ministro continua convencido da sua eficácia, pretendendo economizar nos próximos anos **4 biliões de euros** graças a uma «reforma de Estado», que será apresentada à *troika* em fevereiro e que é muito criticada pela oposição.[20]

O que também vai aparecendo são algumas indecisões curiosas que misturam a escala longa para o euro e a escala curta para o dólar (1 euro vale mais de 1,3 dólares):

> Esta Grande Recessão já custou quase **53 triliões de dólares** (mais de **37 biliões de euros**) e obrigou os estados a desembolsarem 13 triliões de dólares (mais de 9 biliões de euros) no seu combate. O que soma mais de um ano de produto mundial.[21]

Onde, sim, se notam alterações em relação ao que seria uma utilização mais natural da escala longa é em números com casas decimais — «mil milhões» é tratado como se fosse uma unidade inseparável. Esta prática está largamente disseminada na imprensa portuguesa e na DGT:

> Os prejuízos semestrais históricos do BES — **3,57 mil milhões de euros** — revelaram que a almofada de 2000 milhões de euros afinal não era suficiente e atiraram o rácio de solidez para 5%, bem abaixo do mínimo de 7% exigido pelo Banco de Portugal.[22]

Porém, também se encontram exemplos de um uso mais natural da escala longa:

> O BES teve prejuízos de **3577 milhões** nos primeiros seis meses do ano, o que compara com os 237 milhões registados em igual período de 2013.[23]

Mas, tal como o fenómeno da hiperinflação trouxe os grandes números para o dia a dia das populações, também a financeirização da economia e a atual crise bancária e das dívidas públicas traz

com frequência à baila a referência (tradução) a números da ordem de 10^{12}, os nossos biliões «longos» (***trillion*** em inglês, *billions* em francês).

Há que estar também preparado para referir quantidades dessa ordem de grandeza. Como exemplo, em 2001 apenas dois Estados-Membros da UE (Itália e Alemanha) tinham uma **dívida pública bruta** superior a um bilião de euros, mas em 2012 eram quatro (Alemanha, Itália, França, Reino Unido)[(24)], ascendendo, em 2012, a dívida pública bruta dos atuais 28 a 11,03 biliões de euros. No início de 2014, a dívida pública bruta dos Estados Unidos ultrapassava 17,30 biliões de dólares[(25)]. Também se tem assistido nos últimos anos a um aumento significativo das **reservas cambiais** da China, as quais ultrapassam já 3,66 biliões de dólares[(26)]. Por seu lado, os valores do **PIB** de 2013 das dez maiores economias mundiais estão todos eles bem acima do bilião de dólares[(27)], entre 16,72 biliões nos Estados Unidos (1.ª) e 1,67 biliões na Índia (10.ª), passando por 9,33 biliões na China (2.ª) e 2,19 biliões no Brasil (7.ª). O PIB total da União Europeia alcança 16,95 biliões de dólares.

Já a nomenclatura dos grandes números acima de 10^{12} é, essa sim, de utilização muito mais rara. Para informação, apresenta-se a lista dos termos referentes aos números até 10^{63} e os respetivos equivalentes em inglês, francês e português do Brasil.

	pt_{PT}	**en**	**fr**	**pt_{BR}**
10^{15}	mil biliões	quadrillion	billiard	quatrilhões
10^{18}	trilião	quintillion	trillion	quintilhão
10^{21}	mil triliões	sextillion	trilliard	sextilhão
10^{24}	quatrilião	septillion	quadrillion	setilhão
10^{27}	mil quatriliões	octillion	quadrilliard	octilhão
10^{30}	quintilião	nonillion	quintillion	nonilhão
10^{33}	mil quintiliões	decillion	quintilliard	decilhão
10^{36}	sextilião	undecillion	sextillion	undecilhão
10^{39}	mil sextiliões	duodecillion	sextilliard	duodecilhão
10^{42}	septilião	tredecillion	septillion	tredecilhão
10^{45}	mil septiliões	quattuordecillion	septilliard	quatuordecilhão
10^{48}	octilião	quindecillion	octillion	quindecilhão
10^{51}	mil octiliões	sexdecillion	octilliard	sexdecilhão
10^{54}	nonilião	septendecillion	nonillion	septendecilhão
10^{57}	mil noniliões	octodecillion	nonilliard	octodecilhão
10^{60}	decilião	novendecillion	décillion	novendecilhão
10^{63}	mil deciliões	vigintillion	décilliard	vigintilhão

«Modernidade»* vs. *normalização

É significativo o papel que a tradução «colada ao original» pode ter para a mudança da terminologia portuguesa ao longo do tempo, sobretudo se aliada ao desconhecimento de normas e legislação em vigor e à vontade de copiar o que vem de fora e é considerado «mais moderno». O ensino universitário em língua inglesa, corrente em Portugal na área da economia, também vem ajudar a criar instabilidade na terminologia portuguesa. Assim, ao ler textos em português convém saber de que «escola» é o autor (ou tradutor).

Existindo documentos normativos, esses devem ser cumpridos (se não, para que serve a normalização?). Até indicação em contrário (revogação da legislação em vigor e da norma portuguesa NP 18:2006 e consequente revisão do *Código de Redação Interinstitucional*), a DGT, apesar da mudança da língua da maioria dos originais do francês para o inglês, deverá continuar a utilizar a escala longa (regra 6N).

Situação atual noutras línguas

Mas estará Portugal isolado na manutenção da escala longa? Ou, pondo a questão de outra forma, haverá mais países que não utilizem a escala curta?

Entre as 24 línguas oficiais da **União Europeia**[28], verifica-se que a escala longa é utilizada pela maioria das línguas dos países da Europa continental, em conformidade com as recomendações da IX Conferência Geral dos Pesos e Medidas. Por seu lado, a Grécia e os Estados-Membros insulares (Reino Unido, Irlanda, Malta e Chipre) utilizam atualmente a escala curta. A Roménia, a Estónia, a Letónia e a Lituânia utilizam a escala longa para os grandes números mais frequentes (até 10^9 ou 10^{12}) e a escala curta daí para diante.

	10^6	10^9	10^{12}	10^{15}	10^{18}	10^{21}
bg	милион	милиард	билион	билиард	трилион	трилиард
cs	milion	miliarda	bilion	biliarda	trilion	triliarda
da	million	milliard	billion	billiard	trillion	trilliard
de	Million	Milliarde	Billion	Billiarde	Trillion	Trilliarde
el	εκατομμύριο	δισεκατομμύριο	τρισεκατομμύριο	τετράκις εκατομμύριο	πεντάκις εκατομμύριο	εξάκις εκατομμύριο
en	million	billion	trillion	quadrillion	quintillion	sextillion
es	millón	mil millones	billón	mil billones	trillón	mil trillones
et	miljon	miljard	triljon	kvadriljon	kvintiljon	sekstiljon
fi	miljoona	miljardi	biljoona	tuhat biljoonaa	triljoona	tuhat triljoonaa
fr	million	milliard	billion	billiard	trillion	trilliard
ga	milliún	billiún	trilliún	cuaidrilliún	cuintilliún	seisilliún
hr	milijun	milijarda	bilijun	bilijarda	trilijun	trilijarda
hu	millió	milliárd	billió	billiárd	trillió	trilliárd
it	milione	miliardo	bilione	biliardo	trilione	trilliardo
lt	milijonas	milijardas	trilijonas	kvadrilijonas	kvintilijonas	sekstilijonas
lv	miljons	miljards	triljons	kvadriljons	kvintiljons	sekstiljons
mt	miljun	biljun	triljun	kwadriljun	kwintiljun	sistiljun
nl	miljoen	miljard	biljoen	biljard	triljoen	triljard
pl	milion	miliard	bilion	biliard	trylion	tryliard
pt	milhão	mil milhões	bilião	mil biliões	trilião	mil triliões
ro	milion	miliard	mie de miliarde	cvadrilion	cvintilion	sextilion
sk	milión	miliarda	bilión	biliarda	trilión	triliarda
sl	milijon	milijarda	bilijon	bilijarda	trilijon	trilijarda
sv	miljon	miljard	biljon	tusen biljoner	triljon	tusen triljoner
IATE	**796531**	**780587**	**2251016**	**2251017**	**2251018**	**2251019**

Fonte: IATE, http://iate.europa.eu/switchLang.do?success=mainPage&lang=pt.

Mas não é só a maioria dos Estados-Membros que não utiliza a escala curta americana, também uma boa parte da Ásia usa outros sistemas de denominação dos grandes números[29]. Tal é o caso dos países do Extremo Oriente, da Ásia Meridional e do Sudeste Asiático.

Nas línguas do **Extremo Oriente**, tal como na Grécia Antiga[30], a denominação dos grandes números é baseada na **miríade**[31] (dez mil, 10^4). A nomenclatura dos grandes números não se organiza nem em grupos de seis algarismos (escala longa) nem em grupos de três algarismos (escala curta), mas sim em grupos de quatro algarismos — regra 4N.

	chinês			**japonês**		**coreano**		**português europeu**
	simpl.	tradic.	*pīnyīn*	*kanji*	*rōmaji*	hanja	RR	
10^4	万	萬	wàn	万	man	萬	man	dez mil
10^8	亿	億	yì	億	oku	億	eok	cem milhões
10^{12}	兆	兆	zhào	兆	chō	兆	jo	bilião
10^{16}	京	京	jīng	京	kei, kyō	京	gyeong	dez mil biliões
10^{20}	垓	垓	gāi	垓	gai	垓	hae	cem triliões
10^{24}	秭	秭	zǐ	𥝱	shi, jo	秭	ja	quatrilião
10^{28}	穰	穰	ráng	穣	jō	穰	yang	dez mil quatriliões
10^{32}	沟	溝	gōu	溝	kō	溝	gu	cem quintiliões

10^{36}	涧	澗	jiàn	澗	kan	澗	gan	sextilião
10^{40}	正	正	zhēng	正	sei	正	jeong	dez mil sextiliões
10^{44}	载	載	zài	載	sai	載	jae	cem septiliões

Fontes: Wikipedia, *Chinese numerals: Large numbers*, https://en.wikipedia.org/wiki/Chinese_numerals#Large_numbers;
Wikipédia, *Numerais Japoneses*, https://pt.wikipedia.org/wiki/Numerais_japoneses;
Wikipedia, *Korean numerals*, https://en.wikipedia.org/wiki/Korean_numerals.

Esta prática pode ser facilmente observada na versão bilingue chinês-português do *Boletim Oficial da Região Administrativa Especial de Macau*.

> Por exemplo, no Despacho do Chefe do Executivo n.º 403/2008[32] é referida a seguinte quantia em patacas (澳門幣):
>
> - $ 2,213,060,000.00 (澳門幣貳拾貳億壹仟叁佰零陸萬元整)
> - $ 2 213 060 000,00 (dois biliões e duzentos e treze milhões e sessenta mil patacas)
>
> O número 貳拾貳億壹仟叁佰零陸萬 ler-se-á: duas dezenas e duas miríades de miríades (億), um milhar, três centenas, zero e seis miríades (萬). Notar que, pelo menos neste texto, as autoridades de Macau se afastam da terminologia consagrada em Portugal, utilizando a escala curta, escrevendo «dois biliões e duzentos e treze milhões e sessenta mil patacas» em vez de «dois mil duzentos e treze milhões e sessenta mil patacas».

Na **Índia**, a situação é ainda diferente. Até ao milhar, tudo é igual, mas daí para a frente os algarismos agrupam-se aos pares (e não em trios, como é usual nos países ocidentais). A terminologia indiana corresponde a esses agrupamentos.

Na Índia utiliza-se também a escala longa dos tempos do Império Britânico e a escala curta americana, o que, sendo fonte de ambiguidades, favorece objetivamente a manutenção da terminologia tradicional, a qual não se presta a quaisquer confusões.

		hindi	**urdu**	**transcrição**	**português europeu**
10^{3}	1,000	सहस्र	ہزار	sahasra/hazār	mil
10^{5}	1,00,000	लाख	لاکھ	lākh	cem mil/laque
10^{7}	1,00,00,000	करोड़	کروڑ	karoṛ	dez milhões/crore
10^{9}	1,00,00,00,000	अरब	ارب	arab	mil milhões/cem crores
10^{11}	1,00,00,00,00,000	खरब	کھرب	kharab	cem mil milhões/dez mil crores
10^{13}	1,00,00,00,00,00,000	नील	نیل	neel/nīl	dez biliões/dez laques de crores

Fonte: Wikipedia, *Indian Numbering System*, https://en.wikipedia.org/wiki/Indian_Numbering_System

Alguns destes termos chegaram aos textos portugueses na antiga Índia Portuguesa.

> A exportação total foi de 197 **crores** (cada crore tem 100 **laques** ou dez milhões) de rupias, e a importação de 131 crores de rupias, deixando um saldo a favor da Índia de 66 crores.
>
> *Heraldo*, 23 de agosto de 1916[33]

Paulo.Correia@ec.europa.eu

---

(1) Figueiredo, C. de, *Novo Dicionário da Língua Portuguesa*, 1913, http://www.gutenberg.org/files/31552/31552-pdf.pdf.
(2) Correia, P., «Em torno do bilião» *in* «a folha» n.º 18 — primavera de 2005, http://ec.europa.eu/translation/portuguese/magazine/documents/folha18_pt.pdf.
(3) **Regra 3n ou (n-1)**: (n-1)lião = 10^{3n}; exemplo: (n=3) bilião = $10^{3\times3} = 10^{6}$.
(4) **Regra 6N ou N**: (N)lião = 10^{6N}; exemplo: (N=2) bilião = $10^{6\times2} = 10^{12}$.

[5] *Ant*. Mil vezes mil coisas ou pessôas: «concorre a ella tanta gente, que se affirma que passa de tres contos de pessôas». *Peregrinação*, CVIII, Figueiredo, C. de, *Novo Dicionário da Língua Portuguesa*, 1913, http://www.gutenberg.org/files/31552/31552-pdf.pdf.

[6] Alguns exemplos da definição de bilião em dicionários portugueses publicados até 1958:
1868: Lacerda, *Novo Diccionario da Lingua Portugueza* — billião, mil milhões
1873: Roquete, J.-I., *Diccionario da Lingua Portugueza* — billião, mil milhões
1881: Aulete, F. J. C., *Diccionario Contemporaneo da Lingua Portugueza* — billião, mil milhões
1910: de Séguier, J., *Diccionário Prático Illustrado* — billião, mil milhões
1916: Almeida, F. de, *Diccionario Illustrado da Lingua Portugueza* — billião, **um milhão de milhões**
1932: Torrinha, F., *Moderno Dicionário da Língua Portuguesa* — bilião, mil milhões
1958: Machado, J. P., *Dicionário de Língua Portuguesa* — bilião, mil milhões

[7] Em 1922, a nota de maior valor emitida pela República de Weimar (Alemanha) era de 50 mil marcos e, em finais de 1923, era de 100 biliões (100 000 000 000 000) de marcos.

[8] Bureau internationale des poids et mesures, *Comptes rendus des séances de la neuvième Conférence générale des poids et mesures réunie à Paris en 1948*, p. 71, http://www.bipm.org/utils/common/pdf/CGPM9.pdf.

[9] «Décret n.º 61-501 du 3 mai 1961 relatif aux unités de mesure et au contrôle des instruments de mesure», *Journal officiel de la République française*, 20.5.1961, http://www.legifrance.gouv.fr/jopdf/common/jo_pdf.jsp?numJO=0&dateJO=19610520&numTexte=&pageDebut=04584&pageFin=.

[10] Em 1974, Harold Wilson, à época primeiro-ministro britânico, respondeu por escrito a uma pergunta colocada na Câmara dos Comuns: «The word "billion" is now used internationally to mean 1,000 million and it would be confusing if British Ministers were to use it in any other sense.», Wikipedia, *Long and short scales*, http://en.wikipedia.org/wiki/Long_and_short_scales.

[11] «The American system of numeration for denominations above one million was modeled on the French system but more recently the French system has been changed to correspond to the German and British systems», Merriam-Webster, *Webster's Ninth New Collegiate Dictionary*, 1985, p. 810.

[12] «Portaria n.º 17052 que determina que no ensino da numeração falada e nos livros didácticos se adopte na nomenclatura dos grandes números a nomenclatura internacional da regra N», *Diário da República*, n.º 49, I Série, de 4 de março de 1959, http://dre.pt/pdf1sdip/1959/03/04900/02420242.pdf.

[13] Na sequência da revisão da NP 18:1953, aprovada pela «Portaria n.º 14608 que aprova como definitiva, com o n.º NP-18, a norma provisória P-18, relativa a nomenclatura dos grandes números», *Diário da República*, n.º 250, I Série, de 11 de novembro de 1953, http://www.dre.pt/pdf1s/1953/11/25000/13831383.pdf.

[14] Instituto Português da Qualidade, «NP 18:2006 (Ed. 3): Nomenclatura dos grandes números», http://www.ipq.pt/custompage.aspx?modid=0&pagid=1250&TPA=C&ncert=94876.

[15] Alguns dicionários brasileiros registam os galicismos «miliardário» e «milhardário».
Cf. Editora Melhoramentos, *Michaelis Dicionário de Português Online*, http://michaelis.uol.com.br/moderno/portugues/index.php?lingua=portugues-portugues&palavra=miliardário.

[16] Cf. Serviço das Publicações, *Código de Redação Interinstitucional: 10.9.1. Emprego dos algarismos árabes*, http://publications.europa.eu/code/pt/pt-4100901pt.htm.

[17] Direção-Geral da Tradução, *English Style Guide*, 2014, http://ec.europa.eu/translation/english/guidelines/documents/styleguide_english_dgt_en.pdf.

[18] «Corte de 4 mil milhões na despesa não é suficiente», *Diário de Notícias*, 21.12.2012, http://www.dn.pt/inicio/economia/interior.aspx?content_id=2960362.

[19] Reproduz-se uma passagem de artigo publicado originalmente no semanário *Expresso* de 19.1.2008:
A diferença entre o bilião "europeu" e o bilião "americano" tem mexido muitos zeros. É por isso que Daniel Amaral lançou, na sua crónica no *Expresso*, dia 22 de dezembro, um repto para uniformizar a linguagem. O economista explica que no mundo financeiro é muita a confusão que se faz entre os dois biliões. Mesmo com escalas diferentes, Daniel Amaral afirma que o bilião utilizado na economia é o "americano". A opinião é partilhada por João Duque, professor do Instituto Superior de Economia e Gestão, que diz que a confusão entre biliões é um problema que se põe na informação financeira. Para Silva Lopes, presidente do Montepio Geral, as confusões entre o bilião "americano" e o "europeu" terminavam com a uniformização da linguagem. «Nem que tivesse de se inventar outra palavra.» O banqueiro adverte que na economia «um bilião com 9 zeros é um número importante, enquanto o bilião com 12 zeros é pouco utilizado». E que mesmo que esteja a discutir números com outro europeu, o bilião "americano" é o assumido. «Só com os franceses é que dizemos o *milliard*, porque eles arranjaram uma palavra própria.» Se algum dia se avançar para uma alteração da palavra, o português não será um entrave. «A língua é uma unidade mutável e a tendência é para adotar o bilião "americano", até pela própria influência da cultura norte-americana», diz o professor de Português Pinto Amaral. No entanto, até agora ainda não surgiram propostas de adoção da escala norte-americana. A garantia vem de Eduarda Filipe, do Laboratório de Metrologia do Instituto Português da Qualidade. Desde 1948 que o sistema se mantém. Tudo porque a 9.ª Conferência de Pesos e Medidas recomendou a escala longa para os países europeus. O professor Guilherme de Almeida, autor do livro *Sistema Internacional de Unidades, Grandezas e Unidades Físicas, Terminologia, Símbolos e Recomendações*, alerta que uma uniformização da linguagem não pode partir de uma iniciativa nacional. «Se formos os únicos a adotarmos a escala curta, então afastamo-nos da Europa.» E diz que este problema não se põe no mundo científico. «Na ciência damos primazia à numeração, em vez do substantivo.» Já para Carlos Fiolhais, professor de Física na Universidade de Coimbra, existem outras questões mais urgentes para padronizar. «Nos biliões, o importante é que quando se faça a tradução se tenha em conta a escala.»

[20] «Austeridade excessiva – rigor sem precedentes», *Folha de Portugal*, 6 a 12.1.2013, http://folhadeportugal.pt/austeridade-excessiva-rigor-sem-precedentes/.

[(21)] «Crise mundial já "custou" 46 biliões de euros», *Expresso*, 28.2.2010, http://expresso.sapo.pt/crise-mundial-ja-custou-46-bilioes-de-euros=f568122#ixzz3Bl8hpFDo.
[(22)] «Operação relâmpago: BES desaparece e nasce Novo Banco», *Dinheiro Vivo*, 4.8.2014, http://www.dinheirovivo.pt/Mercados/Banca/interior.aspx?content_id=4062395.
[(23)] «Prejuízos do BES ascendem a 3577 milhões de euros», *Expresso*, 30.7.2014, http://expresso.sapo.pt/prejuizos-do-bes-ascendem-a-3577-milhoes-de-euros=f883887#ixzz3BlAaDAea.
[(24)] Fonte: Eurostat, *General government gross debt — annual data: Map*, http://epp.eurostat.ec.europa.eu/tgm/mapToolClosed.do?tab=map&init=1&plugin=1&language=en&pcode=teina225&toolbox=types.
[(25)] US Debt Clock, http://www.usdebtclock.org/.
[(26)] «News: PBOC Says No Longer in China's Interest to Increase Reserves», *Bloomberg*, 21.11.2013, http://www.bloomberg.com/news/2013-11-20/pboc-says-no-longer-in-china-s-favor-to-boost-record-reserves.html.
[(27)] Central Intelligence Agency, *The World Factbook: GDP (official exchange rate)*, https://www.cia.gov/library/publications/the-world-factbook/fields/2195.html.
[(28)] Por iniciativa do departamento de língua portuguesa, passou-se recentemente em revista o conteúdo da base terminológica IATE para os grandes números com o apoio da unidade de coordenação terminológica e dos terminólogos da DGT das atuais 24 línguas oficiais da União Europeia.
[(29)] Wikipédia, *Escalas curta e longa: World map of long and short scales*, https://pt.wikipedia.org/wiki/Escalas_curta_e_longa#mediaviewer/Ficheiro:World_map_of_long_and_short_scales.svg.
[(30)] No grego atual, a miríade subsiste não como palavra autónoma mas apenas como elemento de formação na nomenclatura dos grandes números. Assim, embora dez mil se designe Δέκα χιλιάδες (dez milhares), milhão designa-se εκατομμύριο (algo como hectomiríade, isto é cem miríades).
[(31)] A miríade tem um lugar muito importante na cultura oriental, nomeadamente na expressão **万**岁 / **万歳** — literalmente dez mil anos —, que adquire o sentido da saudação viva!, saúde!, longa vida! A saudação ficou conhecida nos países ocidentais por ter sido amplamente utilizada na China durante a Revolução Cultural (**Longa vida** presidente Mao! / 毛主席**万**岁 / *Máo Zhŭxí* ***wànsuì****!*) e ser corrente no Japão (**Banzai**! / **万**歳), por evolução fonética de *manzai* (*man* = miríade; *zai* = ano).
[(32)] «Despacho do Chefe do Executivo n.º 403/2008», *Boletim Oficial da Região Administrativa Especial de Macau*, n.º 52, I Série, 3.º Suplemento, de 31 de dezembro de 2008, http://images.io.gov.mo/bo/i/2008/52/despce-403-2008.pdf.
[(33)] Dalgado, S. R., *Glossário Luso-Asiático*, Buske, Hamburgo, ISBN 3-87118-479-9, 1.º vol. (A-L), http://books.google.pt/books?id=zOufgbY8TbsC&lpg=PP1&ots=ZOE6VtBkxV&dq=Sebasti%C3%A3o%20Rodolfo%20Dalgado&hl=pt-PT&pg=PP4#v=onepage&q&f=false.

Espanhol e português: as dificuldades inesperadas (II)

Augusto Múrias
Parlamento Europeu

DIFERENÇAS ORTOGRÁFICAS E FONÉTICO-FONOLÓGICAS

Este artigo[(1)] segue-se a uma parte introdutória, publicada no número anterior d'«a folha»[(2)]. Ao procurar ser representativo, não é obviamente exaustivo, nem em quantidade nem em qualidade[(3)]. São exclusivamente as diferenças, e não as coincidências ortográficas, que despertam a nossa atenção, no pressuposto de que uma forma ortográfica equivalente veicula diferenças fonético-fonológicas (p. ex.: *despojos*). Cada vocábulo, limitado a uma determinada aceção, surgirá combinado com o que, na outra língua deste par, apresenta mais afinidades ortográficas e fonético-fonológicas. Porém, uma análise semântica aprofundada não só colocaria em causa o alcance dessa equivalência pressuposta, como acrescentaria outras possíveis correspondências lexicais (como no caso de *permiso/permissão*).

As «tabelas de correspondência» entre ambas as línguas que figuram mais adiante não pretendem sugerir a frequência com que uma dada correspondência ocorre. Há, isso sim, fenómenos casualmente mais exemplificados que outros. Concentramo-nos por vezes num aspeto, ignorando outros que também viria a propósito referir: por exemplo, considerar em *gitano/cigano* uma única consoante diferente, quando na verdade são duas; não ter em conta que uma diferença é condicionada por outro fenómeno, como na vocalização em português *millonario/milionário*; apreciar os vocábulos *feria* (PT:

feira) ou *internado* (PT: internato) enquanto diferença formal, e não também semântica (falso amigo). Na verdade, determinadas diferenças formais poderiam ser antes tratadas numa perspetiva morfológica (caso do prefixo em *encender*) ou semântica (*escupir/cuspir* como falso amigo). Podemos aleatoriamente tomar como exemplo uma forma simples, ignorando as respetivas formas derivadas ou flexionadas, ou vice-versa [(4)].

Sintomático da forte analogia existente de princípio entre ambos os idiomas em análise é a profusão de casos de divergência ortográfica e fonético-fonológica apenas ligeira[(5)]: «*O aluno de nível intermediário e avançado que já possui um conhecimento razoável do espanhol deve estar alerta para, no caso de produção espontânea, não aportuguesar termos como femenino, cemento, soborno, diputado, etc., uma vez que é apenas um fonema que diferencia o vocábulo espanhol do português.*»[(6)] Esta dificuldade é particularmente problemática «*cuando la palabra es pronunciada igual al portugués y tiene grafía distinta, como profesor, pasar y zapato, cabeza (para quienes sesean), o al revés, que se escribe igual y se pronuncia de forma distinta, como casa, gente, (...) cómodo*»[(7)]. Cabe por isso lembrar que «*La pronunciación es un factor esencial en el proceso de aprendizaje de la lengua como parte del componente lingüístico de la competencia comunicativa*»[(8)]. Isto deve-se ao facto de que o grau de aceitação do aluno de uma língua estrangeira «*en la comunidad extranjera será tanto mayor como mejor sea su acento.*»[(9)]

Intuitivamente, o falante nativo do português tende a generalizar certas divergências formais:
— à nasalação final em português, marcada por «m» ou por vogal ou ditongo nasais, pode corresponder em espanhol «n» final (ES: *en común; desdén; afán; pelotón;* PT: em comum; desdém; afã; pelotão; mas: *crimen* (crime), *volumen* (volume)); outras possibilidades em espanhol para a nasalação final em português ilustram as palavras *vano* (vão), *rasguño* (rasgão), *anfitriona* (anfitriã), *frenesí* (frenesim), *maniquí* (manequim), *nube* (nuvem);
— redução sistemática do duplo «ss» em português para um único «s» em espanhol: *desasosiego*; el *asunto* de las *comisiones*; *agresivo; tosir; presentir*[(10)];
— inexistência em espanhol dos grupos nh, lh, ç;
— conversão de *-je* em final de palavra em *-gem* (*chantaje/ chantagem, salvaje/ selvagem*); de *-able* em *-ável* (*razonable/razoável*); de *-ción* em *-ção*; de *-ales* em *-ais* (*perjudiciales/prejudiciais; vendavales/vendavais*); de *-ible* em *-ível* (*ostensible/ostentível*); etc.

Tais generalizações podem nuns casos fazer mais sentido que noutros:
— terminação em *-iendo/-yendo* no caso do gerúndio dos verbos dos grupos *-er* e *-ir* (*comprometiendo, contradiciendo, leyendo, se cayeron, yendo y viniendo*), mas em *-endo* em caso de adjetivo (*horrendo, tremendo*) e de nome (*decrescendo, subtraendo*; mas: *remiendo*).

Além disso, são múltiplas as restrições a outras generalizações:
— *contienda, enmienda*; mas: *calenda, reprimenda*
— *entendimiento, fingimiento, llamamiento, movimiento*; mas: *divertimento, fomento, pavimento, ornamento*
— *aguardiente, corriente, requeriente, serpiente, decendiente, independiente*; mas: *accidente, antecedente, indiferente, reticente.*

No tocante às dificuldades ortográficas, assinalem-se ainda as seguintes:
— a utilização de «h» no início de palavra em espanhol, ao contrário dos respetivos correspondentes em português: *hombro, hierba, huérfanas, hueso, huevo* (mas: *armónica, armonía, armonizar, armoniosamente*). Esta dificuldade é ainda maior quando a letra «h» surge em princípio de sílaba no meio da palavra, possibilidade esta que não existe em português moderno: *adherencia, alcohol, ahí, aprehensión* (aprensión), *bohemio, bonhomía, deshonesto, exhalar, exhibía, exhortación, hueco, inhibir, inhóspito, prohibir, vehemente*. Acresce que um verbo como *«oler» (cheirar)* pode ter formas iniciadas por «h»: *¿Hueles lo que yo huelo?*
— a utilização (mais frequente) de «y» em espanhol (*apoyo, ensayo, influyente, inyección, yantar, suyo, lacayo*)[(11)] e da letra «k» (*folklore, folklórico, kilómetro*).

— a utilização de diéresis para marcar em espanhol a pronúncia de cada elemento constitutivo do grupo *-güe* ou *-güi*: *antigüedad, vergüenza, argüir, pingüino*. Em contrapartida, atente-se na inexistência em espanhol dos acentos grave, til e circunflexo do português: *lámpara/lâmpada*.
— o desvio ortográfico na realização de propriedades fonético-fonológicas amplamente coincidentes em ambas as línguas: *cotidiano* (quotidiano), *cincuenta* (cinquenta), *consecuencia* (consequência)[12]. A forma **qua* não existe em ES: *cuarenta* (quarenta), *cuadrado* (quadrado), *cuando* (quando), *cualquier* (qualquer), *cuartel*, *cualitativa* e *cuantitativamente*, *anticuado*.
—a maior quantidade de nomes terminados em *–is* do que em português, do domínio da medicina *(psiconálisis, meningitis, tuberculosis, esquimosis, parálisis, celulitis)* e não só *(antítesis, apocalipsis, brindis, sintaxis)*; a existência em espanhol, ao contrário do português, de vocábulos terminados em *-ea* ou *–eo*, etc.

Diferenças ortográficas implicam, regra geral, diferenças fonético-fonológicas. O novo acordo ortográfico do português, ao prever nomeadamente a queda das consoantes não pronunciadas, associa visualmente diferenças ortográficas entre ambos os idiomas às diferenças fonético-fonológicos que já existiam de antemão: *tacto, práctico, atractivo, acción, adoptar*, etc. Cf. em contrapartida *mentecato* (PT: *mentecapto*).

Iniciaremos a tabela que se segue aludindo aos erros ortográficos que se devem ao facto de uma palavra se escrever junto ou separado. Tomemos como exemplo *por qué* como locução interrogativa (*¿Por qué no viniste ayer a la fiesta?*); *porqué* como nome (*No entiendo el porqué de lo que me dices)*; *porque* como conjunção causal (*No fui a la fiesta porque no tenía ganas*); e *por que+nome* (*Este es el motivo por (el) que te llamé; ¡Por qué calles más bonitas pasamos!*).

Em espanhol, ao contrário do português, pode escrever-se junto certas formas pronominais: (infinitivo) *Me resta el consuelo de haberlo obligado (de o ter obrigado)*. Já pelo contrário, muitas das contrações existentes em português não são permitidas em espanhol: *después de aquello (daquilo;* mas existe a contração *del*: de+el/*do*); un ángel *de la (da)* guarda especial que no se separa *de él (dele)*; *de aquel (daquele)* viaje, la vida *de aquella (daquela)*; gastar *de más (demais)* con el anillo. No me puede llevar *a la (à)* fuerza (mas existe a contração *al*: a+el); El gobierno gastará 2,5 millones *en una (numa)* campaña.

Esperamos ter assim ilustrado a relação de falsa analogia de princípio entre ambos os idiomas no domínio aqui em apreço. No próximo artigo, abordaremos então as diferenças morfológicas e gramaticais.

augusto.murias@europarl.europa.eu

Espanhol	**Português**
Uma palavra/Palavras separadas	
uma palavra	**palavras separadas**
altibajos, donjuan, Jesucristo, Latinoamérica veinticuatro, veintiún, anoche, anteanoche anglosajón, bienestar, bajorelieve, contraataque exmandatario, extremaunción, manirroto, mediodía padrenuestro, pisapapeles, portavoz, puercoespín, sinfín un sinvergüenza, subrepticio, bienintencionado hispanoamericana, malhumorado, pequeñoburguesa el uno enfrente del otro, enseguida, aparte Nochebuena	altos e baixos, Dom[13] Juan, Jesus Cristo, América Latina vinte e quatro, vinte e um, ontem/anteontem à noite anglo-saxão, bem-estar, baixo-relevo, contra-ataque ex-mandatário, extrema-unção, mãos-rotas, meio-dia pai-nosso, pisa-papéis, porta-voz, porco-espinho, sem-fim sem-vergonha, sub-reptício, bem-intencionado hispano-americana, mal-humorado, pequeno-burguesa um em frente ao outro, em seguida, além de noite da consoada
palavras separadas	**uma palavra**
a pesar, a través, en fin, tal vez de prisa, entre dientes, sobre todo, materia prima	apesar, através, enfim, talvez depressa, entredentes, sobretudo, matéria-prima
Outras divergências formais	
3 vogais poseía **2 vogais** correo, creó, europea, atea, ansío, odian	**3 vogais**[14] possuía **3 vogais** correio(s), criou, europeia, ateia, anseio, odeiam

3 vogais	limpiáis	**2 vogais**	limpais
2 vogais	atraer, saraos, reina, compadreo, murmureo	**2 vogais**	bebeu, encolheu, rainha, justiceiro, atenuou, atrair
	trofeo, reo, justiciero, bebió, encogió, héroe, buey		murmúrio, troféu, réu, compadrio, herói, boi
	jueces, malagüero, atenuó, mortuoria		juízes, agoiro, mortuária, saraus
2 vogais	reír, maestro	**1 vogal**	rir, mestre
Ditongo		**Vogal**	
AI	aire, aireada, (conj. irreg.) traigo	**A**	ar, areada, trago
AI	aislado	**I**	isolado
AU	cautivar, cautiverio	**A**	cativar, cativeiro
EA	trajeado, cuchichear, masejear, mordisquearse	**A**	trajado, cochichar, massajar, mordiscar
	pedalear		pedalar
EA	chimenea, marea	**E**	chaminé, maré
EI	veinte, treintena	**I**	vinte, trintena
EO	boxeo	**E**	boxe
EO	cuchicheo, gaseosa	**O**	cochicho, gasosa
IA	bestia, envidia, guardia, herencia, menopausia	**A**	besta, inveja, guarda, herança, menopausa
	salvaguardia, soberbia, Alemania, Francia		salvaguarda, soberba, Alemanha, França
	exiliado, desperdiciar, limpiar, telegrafiar		exilado, desperdiçar, limpar, telegrafar
	(*mas:* enturbiar)		(enturvecer)
IE	ingeniería[15]	**A**	engenharia
IE[16]	apariencia, conocimiento, Estados Miembros, fiera	**E**	aparência, conhecimento, Estados-Membros, fera
	incienso, infierno, limpieza, pariente, piel		incenso, inferno, limpeza, parente, pele
	presentimiento, vientre, ciego, cubierto, prieto		pressentimento, ventre, cego, coberto, preto
	recién, siniestro, tierno, quiere, sugiere		recém, sinistro, terno, quer, sugere
	comienza, pienso, dieron, recomiendo, tiembla		começa, penso, deram, recomendo, treme
IE	siguiente, puntapié, miedo, arriesgarse, diezmar	**I**	seguinte, pontapé, medo, arriscar-se, dizimar
	(conj. irreg.) me divierto, hiriese (cf. distrajese)		divirto-me, ferisse (distraísse)
IU	ciudad, ciudadano	**I**	cidade, cidadão
IO	agrio	**E**	acre
IO	agravio, aprecio, espacio, vidrio, estudios	**O**	agravo, apreço, espaço, vidro, estudos
	limpio, sucio, amplio, nervioso		limpo, sujo, amplo, nervoso
	garfio, maniobra		garfo, manobra
UA	cuaderno, mensualidad, menguado	**A**	caderno, mensalidade, mingado
UE[17]	anzuelo, cuento, estruendo, huésped, muestra	**O**	anzol, conto, estrondo, hóspede, mostra
	un buen lote, cigüeña, luego, risueño		bom, cegonha, logo, risonho
	supuestamente, (conj. irreg.) cuentan, duermes		supostamente, contam, dormes
	puesto (*mas:* cuero)		posto (couro)
UE	trueque	**U**	truque
UI	buitre	**U**	abutre
UO	majestuoso, antiguo, monstruo	**O**	majestoso, antigo, monstro
Vogal		**Ditongo**	
A	faja, más	**AI**	faixa, mais
A	ingesta, jerga	**ÃO**	ingestão, jargão
A	basado, peinado, vagar	**EA**	baseado, penteado, vaguear
A	somera	**IA**	sumária
A	neurona	**IO**	neurónio
A	flotaba, garita	**UA**	flutuava, guarita
E	barbería, mercería	**EA**	barbearia, mercearia
E	astillero, cadera, madeja, queso, sujeto, trasero	**EI**	estaleiro, cadeira, madeixa, queijo, sujeito, traseiro
	parapeto, pecho, solterona, besar, enterarse		parapeito, peito, solteirona, beijar, inteirar-se
	lego, falta de respeto (*mas:* respecto a la crísis)		leigo, respeito
I	golosinas, (conj. irreg.) me di cuenta de	**EI**	guloseimas, dei-me conta de
I	mi	**EU**	meu
O	pavo, permiso	**ÃO**	pavão, permissão
O	fisonomía, multimillonario	**IO**	fisionomia, multimilionário
O	(conj. irreg.) introdujo,produjo	**IU**	introduziu, produziu
O	azote, faros	**OI**	açoite, faróis
O	o[18], otros, dorado, loable, moco, osadía	**OU**	ou, outros, dourado, louvável, mouco, ousadia
	ropa, tocino, otorgar, reposo, oír		roupa, toucinho, outorgar, repouso, ouvir
O	lujoso	**UO**	luxuoso
U	su	**EU**	seu
U	(conj. irreg.) supiste	**OU**	soubeste
U	distribución	**UI**	distribuição
Vogal dupla		**Vogal simples**	
Usted cree que… leer (*mas:* poseer)		crê, lê (possuir)	
Vogal simples		**Vogal dupla**	

comprender, emprendedor, reprensión		compreender, empreendedor (empresário), repreensão	
Vogal diferente[19]			
A	alarma, asamblea, barniz, estrofa, pagoda, sirena	**E (átono)**	alarme, assembleia, verniz, estrofe, pagode, sirene
	resplandor, adinerado, agotado, asequible		resplendor, endinheirado, esgotado, exequível
	piadosa, salvaje, atardecer, aguantar		piedosa, selvagem, entardecer, aguentar
	apiadarse		apiedar-se
E	alarma, asesinato, desventaja, tapicería	**A**	alarme, assassinato, desvantagem, tapeçaria
	trenza, compinches, rencor, mermelada		trança, compinchas, rancor, marmelada
	averiado		avariado[20]
E	derechos, edad, enemigo, gravedad, hipocresía	**I**	direitos, idade, inimigo, gravidade, hipocrisia
	laberinto, lengua, vecino, venganza		labirinto, língua, vizinho, vingança
	creativa, demoledor, exento, femenino, húmeda		criativa, demolidor, isento, feminino, húmida
	peor, profesionales, eso, esto, corregir, toser		pior, profissionais, isso, isto, corrigir, tossir
E	goles, relieve, amedrentar, valeroso, de bruces	**O**	golos, relevo, amedrontar, valoroso, de bruços
I	alquitrán, linterna, bachirillato	**A**	alcatrão, lanterna, bacharelato
I	automóvil, avaricia, canilla, difunto, dimisión	**E**	automóvel, avareza, canela, defunto, demissão
	ingenio, interviniente, liviandad, martillo, sillín		engenho, interveniente, leviandade, martelo, selim
	varilla, víspera, disgustado, efímera, mismo		vareta, véspera, desgostado, efémera, mesmo
	anticiparse, conocimos, concibes, convertir		antecipar-se, conhecemos, concebes, converter
	disfrutar, gimió, irguiendo, sintió, sugirió		desfrutar, gemeu, erguendo, sentiu, sugeriu
I	saltimbanqui	**O**	saltimbanco
I	tiburón, derribar	**U**	tubarão, derrubar
O	roñoso, rescoldo	**A**	ranhoso, rescaldo
	(pret. perf. simples, 3.ª pes. pl.) encontraron		encontraram
O	apetito, micrófono, reclamo, trasto, a oscuras	**E**	apetite, microfone, reclame, traste, às escuras
	a rebato		a rebate
O	corta, encorvada, europeos, sofocante, sordo	**U**	curta, encurvada, europeus, sufocante, surdo
	todo, costar, embrollar, sorprender		tudo, custar, embrulhar, surpreender
	orinar, soster, manosear, sospechó		urinar, suster, manusear, suspeitou
	aplomo, complicidad, fondo, momia, ombligo		aprumo, cumplicidade, fundo, múmia, umbigo
	ortigas		urtigas
U	(conj. irreg.) me detuve	**I**	detive-me
U	carrusel, gusto, puntapié, rulo, suma	**O**	carrossel, gosto, pontapé, rolo, soma
	puntual, tumbado, interrumpir, burdel, cubil		pontual, tombado, interromper, bordel, covil
	descubridor, dulzura, espíritu, ímpetu, ocurrencia		descobridor, doçura, espírito, ímpeto, ocorrência
	tribu, concurrido, riguroso, cubrir, jugar, sufrir		tribo, concorrido, rigoroso, cobrir, jogar, sofrer
	durmiendo (dormir)		dormindo (dormir)
Consoante		**vogal (vocalização)**	
cadena, doctrina, efecto, estreno, pasillo, tabla, recelo, velo		cadeia, doutrina, efeito, estreia, passeio, tábua, receio, véu	
retraso, claraboya, secta, enronquecido, perfecto, yodo		atraso, claraboia, seita, enrouquecido, perfeito, iodo	
ensangrentada, aceptar, pernoctar, valdría, jerarquía, jerga		ensanguentada, aceitar, pernoitar, valeria, hierarquia, jargão	
Vogal		**Consoante (consonantização)**	
absolución		absolvição	
Consoante+vogal		**Vogal**	
los quehaceres, retrasado		afazeres, atrasado	
Consoante sonora		**Consoante surda (ensurdecimento)**	
B	alcoba	**F**	alcofa
D	estadística, internado	**T**	estatística, internato
G	agrio	**C**	acre
G	la consiguiente devaluación	**QU**	consequente
G (intervocálico)	haga, oiga	**Ç**	faça, oiça
G (intervocálico)	encogido, recogimiento[21]	**LH**	encolhido, recolhimento
J (intervocálico)[22]	herejía, pájaro, tejer, dijeron, jabón	**(S)S/C/Ç**	heresia, pássaro, tecer, disseram, sabão
	cobijar		cobiçar
J (intervocálico)		**X** (diferentes valores fonético-fonológicos); **CH**	
	bruja, lujo, ojalá, cajón, mejillones		bruxa, luxo, oxalá, caixote, mexilhões
	quejarse, pujar, vejación, cojo, rajado		queixar-se, puxar, vexação, coxo, rachado
	complexo, prolijas; ejemplo, ejercicio		complexo, prolixas; exemplo, exercício
LL	llamas, lleno, llorar	**CH**	chamas, cheio, chorar
PL	plomo	**CH**	chumbo
Z (vogal)[23]	cazar, izar, lanzar, reforzar, trazar	**Ç/S/C** (vogal)[24]	caçar, içar, lançar, reforçar, traçar
	tropezar, amenaza, azúcar, brazo, cabeza, disfraz		tropeçar , ameaça, açúcar, braço, cabeça, disfarce
	esperanza, hazaña, lodazal, mazmorra, raza		esperança, façanha, lodaçal, masmorra, raça
	regazo, zanahoria, zapatero, postizo		regaço, cenoura, sapateiro, postiço
Z	cabizbajo, izquierdo, lápiz, zafarse	**S**	cabisbaixo, esquerdo, lápis, safar-se
Consoante surda		**Consoante sonora (sonorização)**	

C	masticar, carcajada, secreto, rescatar	**G**	mastigar, gargalhada, segredo, resgatar
	(*mas:* arcilla, sucio)		(argila, sujo)
C	cebra, celo, cocina, docena, estupideces, gacela	**Z**	zebra, zelo, cozinha, dúzia, estupidezes, gazela
	vacío, once, quince, broncear, decir, apaciguar		vazio, onze, quinze, bronzear, dizer, apaziguar
CH[25]	chaqueta, desecho	**J**	jaqueta, desejo
CH	lucha	**T**	luta
J	badajo	**L**	badalo
J	barajar, envejecer, hojear, mejorar, mojar	**LH**	baralhar, envelhecer, folhear, melhorar, molhar
	abeja, agujón, burbuja, espejo, hijo, migaja		abelha, agulhão, borbulha, espelho, filho, migalha
	navaja, ojos, ojeras, paja, piojo , tejado, trabajo		navalha, olhos, olheiras, piolho, telhado, trabalho
	ajeno, semejante		alheio, semelhante
J	garaje, furajido, chantaje, hereje	**G**	garagem, foragido, chantagem, herege
Y	gema, yeso, leyenda, trayectoria	**G**	gema, gesso, legenda, trajetória
Y	conyunctura, inyectar	**J**	conjuntura, injetar
J	empujar	**R(R)**[26]	empurrar
J	enjambre, enjaguarse, prójimo, prolijo	**X**	enxame, enxaguar-se, próximo, prolixo
J	(conj. verb.) produjo	**Z**	produziu
P	apertura, patata, supiste	**B**	abertura, batata, soubeste
S	bisutería, besucar, requesón	**J**	bijutaria, beijocar, requeijão
T	caminata, datos, dificultad, fatiga	**D**	caminhada, dados, dificuldade, fadiga
	deslealtad, pubertad, arrepentirse, bigotito		deslealdade, puberdade, arrepender-se, bigodito
Consoante diferente			
B	árbol, abuelo, barniz, bellaquería, bulto, caballete	**V**[27]	árvore, avô, verniz, velhacaria, vulto, cavalete
	cebada, embestida, librería, haba, sobaco		cevada, investida, livraria, fava, sovaco
	abovedado, estorbar, librarse de, debe, paseaba		abobadado, estorvar, livrar-se, deve, passeava
V	movilización, inmovilizado, vejiga, reventar	**B**	mobilização, imobilizado, bexiga, rebentar
C	ecuestre	**QU**	equestre
CH	archipiélago	**QU**	arquipélago
Q	quirúrgica	**C**	cirúrgica
D	codicia	**B**	cobiça
D	huida	**G**	fuga
D	nido	**NH**	ninho
D	cofradía	**R**	confraria
D	aullido, viudo	**V**	uivo, viúvo
H (início sílaba)	ahogarse, ahorcarse, hervir, hundir	**F**	afogar-se, enforcar-se, ferver, fundir
	ahorro, almohada, bienhechor, deshecha, hastío		aforro, almofada, benfeitor, desfeita, fastio
	higuera, hollín, herido, hormigueo, ahijada		figueira, fuligem, ferido, formigueiro, afilhada
	moho, herramienta, hierro, hondo, humo, rehén		mofo, ferramenta, ferro, fundo, fumo, refém
	(*mas:* ahora, helado, hinchar e rehusar)		(agora, gelado, inchar, recusar)
J	jabón	**S**	sabão
J	crucifijo, prójimo, enjambre, lija, lagartija	**X**	crucifixo, próximo, enxame, lixa, lagartixa
L	pila, silueta	**LH**	pilha, silhueta
L	naranja, alfiler, comulgar	**N**	laranja, alfinete, comungar
L	elementales, mármol	**R**	elementares, mármore
L	coliflor	**V**	couve-flor
N	ganar, pergamino	**NH**	ganhar, pergaminho
Ñ	pestañas	**N**	pestanas
R	surco, franela	**L**	sulco, flanela
S	masónico	**Ç**	maçónico
S	parsimonia	**C**	parcimónia
S	ensartar	**X**	enxertar
T	alvorotar	**Ç**	alvoroçar
X	excavar, extender	**S**	escavar, estender (*mas*: extensão)
Y	yema, cónyuges, yantar, proyector	**J/G+e**	gema, cônjuges, jantar, projetor
	ahuyentar, ayudar, yacer, trayecto, leyenda		afugentar, ajudar, jazer, trajeto, legenda
Y	payaso	**LH**	palhaço
Z	mazmorra	**S**	masmorra
Duas consoantes		**Duas consoantes**	
LH	llagas, hallar	**CH**	chagas, achar
NR	sonreír	**RR**	sorrir
PL	aplomo	**PR**	aprumo
ZG	juzgar (*mas:* sojuzgar, mecenazgo)	**LG**	julgar (subjugar, mecenato)
Consoante dupla (iguais)		**Duas consoantes (diferentes)**	
LL	fallar, humillación, tullido	**LH**	falhar, humilhação, tolhido
LL	llano	**PL**	plano (adj.)
LL	llanto	**PR**	pranto

Consoante dupla (iguais)		**Consoante simples**	
CC	acceder, accidente, diccionario, inspección	**C/Ç**	aceder, acidente, dicionário, inspeção
	satisfacción, fláccido, occidente		satisfação, flácido, ocidente
LL	ampolla, belleza, bullicio, collar, estrella, fuelle	**L**	ampola, beleza, bulício, colar, estrela, fole
	querella, sello, sollozo, aquellos, callado, collar		querela, selo, soluço, aqueles, calado, colar
	escabullir, embellecer, llevar, allí		escapulir, embelezar, levar, ali
LL	murmullo	**R**	murmúrio
NN	innato, innecesario, innegable, innumerable	**N**	inato, desnecessário, inegável, inumerável
	connivencia, ennegrecer, ennoblecer, innovación		conivência, enegrecer, enobrecer, inovação
	perenne		perene
RR	parroquia	**R**	paróquia
Consoante simples		**Consoante dupla (iguais)**	
R	arañar	**RR**	arranhar
S	misionero, sacrosanto	**SS**	missionário, sacrossanto
Duas consoantes (diferentes)		**Consoante simples**	
MN	calumnia, columna, solemnidad, condemner	**N**	calúnia, coluna, solenidade, condenar
	gimnasta, insomnia, somnoliento		ginasta, insónia, sonolento
NM	inmaterial, inmenso, inmerecido	**M**	imaterial, imenso, imerecido
	(n) inmueble, (adj) inmóvil		imóvel (n + adj)
	conmigo, conmevedor, inmoral		comigo, comevedor, imoral
	conmutable, inmediata		comutável, imediata
	conmiseración, inminencia, inmundicia		comiseração, iminência, imundice
	conmemoración, enmienda		comemoração, emenda
	inmigrante, enmudecer, enmarañar, inmicuirse		imigrante, emudecer, emaranhar, imiscuir-se
SC	descifrar	**C**	decifrar
ZC	agradezco, padezco, parezco	**Ç**	agradeço, padeço, pareço
	conozca, desaparezca, permanezca		conheça, desapareça, permaneça
ZG	noviazgo	**D**	noivado
ZG	mecenazgo	**T**	mecenato
CH	derecha, techo	**T**	direita, teto
Consoante simples		**Duas consoantes (diferentes)**	
C	nacimiento, necio, crecer	**SC**	nascimento, néscio, crescer
	conciencia, concienciarse (*mas:* inconsciente)		consciência, consciencializar-se
	acrecentar, convalecer, florecer		acrescentar, convalescer, florescer
N	marinero, ranura	**NH**	marinheiro, ranhura
Ñ	leño	**NH**	lenha
Acréscimo de vogal[28]		**Supressão de vogal**	
INÍCIO DA PALAVRA			
	acreedor, adeudar, ajuedrez, ametralladora, arruga		credor, dever dinheiro, xadrez, metralhadora, ruga
	trae acá, eslógan, esnobismo, obispo, Oporto		cá, slôgane, snobismo, bispo, Porto
MEIO DA PALAVRA			
	bereber, heredar, pertenecer, vanagloriarse		berbere, herdar, pertencer, vangloriar-se
FIM DA PALAVRA			
	aire, ni siquiera, (conj. verbal) supuse		ar, nem sequer, supus
Supressão de vogal		**Acréscimo de vogal**	
INÍCIO DA PALAVRA			
	mañana, palpar, perfeccionar, pitar, presentar		amanhã, apalpar, aperfeiçoar, apitar, apresentar
	pregonar, profundizar, pudrirse/podrirse		apregoar, aprofundar, apodrecer-se
	tirarse, señalar, bóveda, campamento		atirar-se, assinalar, abóbada, acampamento
	redondeado		arredondado
MEIO DA PALAVRA			
	escondrijo, fútbol, perla, pulgar, reglamento		esconderijo, futebol, pérola, polegar, regulamento
	vocablo, sabroso, habría, podrían		vocábulo, saboroso, haveria, poderiam
	disgregarse, entablar, establecer, ofrecer		desagregar-se, entabular, estabelecer, oferecer
FIM DA PALAVRA			
	apóstol, cráter, delgadez, desliz, zinc		apóstolo, cratera, delgadeza, zinco
Inversão da posição de duas vogais			
	ingeniero, vidriera		engenheiro, vidreira
PÓS-POSIÇÃO		ANTEPOSIÇÃO	
	barrio, rabia, ferias, novios, andamios		bairro, raiva, feira, noivos, andaimes
Inversão consoante/vogal		**Vogal/consoante**	
	saludable, apretada, preguntar, disfrazarse		saudável, apertada, perguntar, disfarçar-se
Inversão vogal/consoante		**Consoante/vogal**	
	cáncer, corbata, guirnalda		cancro, gravata, grinalda
Acréscimo de consoante		**Supressão de consoante**	

INÍCIO DA PALAVRA	
donde, despavorido, brisca, retraso, dilucidar	onde, espavorido, bisca, atraso, elucidar
pintoresco	pitoresco
MEIO DA PALAVRA	
adelantar, afilar, almorzar, disminuir, entonar	adiantar, afiar, almoçar, diminuir, entoar
lamer, salir, vigilar, actitud, águila, atraco, celda	lamber, sair, vigiar, atitude, águia, ataque, cela
cansancio, cementerio, conejo, crudeza, defensa	cansaço, cemitério, coelho, crueza, defesa
descripción, establo, langosta, molino, rumbo	descrição, estábulo, lagosta, moinho, rumo
salud, septiembre, octubre, anecdótico	saúde, setembro, outubro, anedótico
despreciativo, dulce, sudado vanidoso	depreciativo, doce, suado, vaidoso
FIM DA PALAVRA	
centenar, arrastrar, peligro, registro, rostro	centena, arrastar, perigo, registo, rosto
cable, diablo, hipótesis, brindis, mano	cabo, diabo, hipótese, brinde, mão
a merced del juez, error	mercê, erro
Supressão de consoante	**Acréscimo de consoante**
PRINCÍPIO DA PALAVRA	
abogado, oscuro, escote, sumergir	advogado, obscuro, decote, submergir
neumático, sumisión, zarista	pneu, submissão, czarista
MEIO DA PALAVRA	
cirujano, deporte, oído, peatón, solfeo	cirurgião, desporto, ouvido, peão, solfejo
pronóstico, propiedad, demostrar, desee, veo	prognóstico, propriedade, demonstrar, deseje, vejo
maullar, rumiar, creíble, propia, salado, soso	miar, ruminar, credível, própria, salgado, sonso
resucitado, simplemente	ressuscitado, simplesmente
FIM DA PALAVRA	
apedrear, orquestar, pestañear, (imperativo) !Di!	apedrejar, orquestrar, pestanejar, Diz!
termo, arrabales, culata, simple	termos, arrabaldes, culatra, simples
Permuta de consoante/sílaba	
cocodrilo, caparazón, apretura	crocodilo, carapaça, abertura
Acréscimo de sílaba	**Supressão de sílaba**
INÍCIO DA PALAVRA	
escalofrío, escena, escenario, escéptico, estanque	calafrio, cena, cenário, cético, tanque
adjunto, en su derredor, por encima de	junto, ao seu redor (à sua volta), por cima de
escupiendo, denegar, ensuciarse, entrenar	cuspindo, negar, sujar-se, treinar
MEIO DA PALAVRA	
cenicero, color, contenedor, culebra, aldeana	cinzeiro, cor, contentor, cobra, aldeã
alrededores, comparecencia, cristiano, dentellada	arredores, comparência, cristão, dentada
escalera, feligrés, generación, moneda, murciélago	escada, freguês, geração, moeda, morcego
naturaleza, orfelinato, orilla, panadería	natureza, orfanato, orla, padaria
palmetadas, pérdida, velatorio, bendecido	palmadas, perda, velório, benzido
ceniciento, caliente, colorado, sanguinolento	cinzento, quente, corado, sangrento
entretenido, calmosamente, apresurar, saludar	entretido, calmamente, apressar, saudar
a mediados de, delante de, en general, en medio de	em meados de, diante de, em geral, no meio de
FIM DA PALAVRA	
anillo, búsqueda, cáscara, lana, riñón, risotada	anel, busca, casca, lã, rim, risota
ermitaño, término, tono, contener, desencadenar	ermita, termo, tom, conter, desencadear
mantener, provenir, crudo, nudo, de mediana edad	manter, provir, cru, nu, de meia idade
impresión vívida	viva
Supressão de sílaba	**Acréscimo de sílaba**
INÍCIO DA PALAVRA	
fregar, palidecer, prestar, señalar	esfregar, empalidecer, emprestar, assinalar
vaciar, prensa	esvaziar, imprensa
MEIO DA PALAVRA	
aburrimiento, calavera, desavenencia, deuda	aborrecimento, caveira, desavença, dívida
párrafo, población, siglo, tortuga	parágrafo, população, século, tartaruga
cargado, liado, dudar (conj. irreg), querrá	carregado, ligado, duvidar, quererá
FIM DA PALAVRA	
cine, ciprés, codo, compás, condena	cinema, cipreste, cotovelo, compasso, condenação
desalojo, desarollo, disimulo	desalojamento, desenvolvimento, dissimulação
hacha, haz de luz, pérdida, reloj, testigo	machado, feixe, perda, relógio, testemunha
carta autógrafa, incluso	autografada, inclusive
aburrirse, inflar, impulsar	aborrecer-se, inflacionar, impulsionar
progresar, supervisar, valorar	progredir , supervisionar, valorizar
Formalmente menos extenso	**Formalmente mais extenso**[29]
tutear (*mas:* tratar de Usted)	tratar por tu
cientos de veces, *miles* de personas	centenas, milhares
el púgil	pugilista

Uno desconfía de todo, de las personas, de los objectos, de las paredes.…son los pequeños los que lo hacen sentirse *a uno* miserable. un viejo *dicho*[30]	uma pessoa ditado
Acentuação	
Quando duas vogais não formam ditongo, emprega-se em espanhol o acento para marcar a vogal da sílaba tónica: *continúe, transeúntes, fuí, frío.* Também nos exemplos seguintes a sílaba tónica é a mesma em ambas as línguas, mas as palavras levam acento em espanhol porque terminam em vogal e são agudas: *aquí, allí, escribí, respondí, París.* Os monossílabos não levam habitualmente acento: *pierdo la fe.* Se em português é o acento agudo que permite distinguir, por exemplo, a forma da primeira pessoa do plural dos verbos com infinito em -ar no presente e no pretérito perfeito simples[31], já em espanhol o mesmo acento permite distinguir monossílabos com diferentes funções/significados: (função interrogativa) *No sé qué* hacer contigo. cf: *me dijo que* me fuera Em espanhol, todos os pronomes interrogativos levam acento: Detallar brevemente *cómo* y *cuándo* se publicaron estos cuentos. cf. … y como no me apetecía hablar… Tú – Él: Y *él, el* jefe ¿qué tal los trata? mí (pessoal) vs. mi (possessivo) más (comparativo/advérbio de quantidade) *vs.* mas (conjunção) Os advérbios de modo terminados em –mente guardam o acento do adjetivo de que derivam: herméticamente, fácilmente, momentáneamente, retóricamente, ávidamente, tímidamente. Devido às regras que regem a junção do pronome pessoal a certas formas verbais ou a formação do plural, emprega-se em espanhol o acento quando a sílaba tónica passa então a ocupar a antepenúltima posição: *Entiéndeme, escúcheme, perdóneme, cállese, siéntese, jóvenes.* Por razões de economia formal, em espanhol, como em português, o acento pode servir para distinguir: (aceções semânticas) desde *hacía* seis meses (temporal) // viéndole ir *hacia* la puerta (local) (adj. demonstrativo) A esta mujer la miro // (pron. demonstrativo) A ésta la toco.	
Vocábulos heterotónicos[32]	
Sílaba tónica em posição distinta	
demo**cra**cia, aristó**tó**crata, poli**cí**a, **cón**clave **lí**mite, an**éc**dota (mas: anec**dó**tico), ce**re**bro at**mós**fera, pa**rá**sito, gra**mó**fono, po**lí**glota, ca**ní**bal im**bé**cil, po**lí**glota	democra**ci**a, aristo**cra**ta, po**lí**cia, con**cla**ve li**mi**te, ane**do**ta (ane**dó**tico), cérebro atmos**fe**ra, para**si**ta, gramo**fo**ne, poli**glo**ta, cani**bal** imbe**cil**, poli**glo**ta
Dificuldades de pronúncia devido a diferenças ortográficas	
Atenuação da consoante final por não ser seguida de vogal em espanhol[33]:	
D contrariedad, formalidad, red **S** tos, interés	contrariedade, formalidade, rede tosse, interesse
Pontuação	
Delimita-se com dois pontos de interrogação/exclamação invertidos apenas o segmento implicado: Estás todo el día ahí sentado sin pegar ni golpe, *¿te pasa algo?, ¿estás enfermo?*	
Palavras estrangeiras	
A adaptação à língua importadora pode revelar diferenças que podem ir a ponto de um estrangeirismo perder a aparência de forma estranha numa das línguas em causa:	
atelier, champán, champú, chequeo, kiosco coctel/cóctel, récord, suéter, córner, coñac clínex, ujier, hamburguesa, frac, capó plató, vermú	ateliê, champanhe, champô, *check-up*,quiosque *cocktail*, recorde, *sweater*, canto, conhaque (lenço do papel), porteiro, hambúrguer, fraque, capô *plateau*, vermute

---

[1] Agradeço à colega tradutora Amparo Fernandez Garcia, que zelou pela correção ortográfica de muitos exemplos em espanhol.

[2] Múrias, M., «Espanhol e português: as dificuldades inesperadas (I)» *in* «a folha», n.º 44 — primavera de 2014, http://ec.europa.eu/translation/portuguese/magazine/documents/folha44_pt.pdf.

[3] Para um estudo mais aprofundado das diferenças aqui em causa, ver, por exemplo, Sandes, E. I. de A., *Análise das dificuldades dos estudantes brasileiros de E/LE na percepção e na produção dos sons aproximantes e nasais em língua espanhola* [tese de doutorado], São Paulo, 2010, http://www.teses.usp.br/teses/disponiveis/8/8145/tde-14102010-091300/pt-br.php;
numa abordagem diferente, Oliveira, A. M. R., «Los alumnos hablan, pero ¿cómo, cuando y qué corregirles?», *V Encuentro brasileño de profesores de español*, Belo Horizonte, 2 a 4.10.2009, *Suplementos marcoELE*, n.° 9, 2009, ISSN 1885-2211, http://marcoele.com/descargas/enbrape/ramos_fonetica.pdf.

[4] Os verbos irregulares constituem amiúde exemplo de complexas diferenças formais, como *pierdo/perco* ou *caigo/caio*.
[5] Para uma exemplificação quantificada destes tipos de erro, ver Lozano Antolín, J. G., «Análisis Contrastivo, Análisis de Errores e Interlengua como Instrumentos para la Enseñanza — Aprendizaje del Español a Lusohablantes Brasileños», *Comunidad Todoele*, p. 23 e sgs., http://todoelecomunidad.ning.com/group/elebrasil/forum/topics/ensenar-una-lengua-proxima?xg_source=activity.
[6] Camorlinga, R., «A distância da proximidade: A dificuldade de aprender uma língua fácil», http://revistas.pucsp.br/index.php/intercambio/article/view/4098.
[7] Takeuchi, N. N., «La semejanza con la lengua materna: tropiezos para el aprendizaje del español», *Revista Letras*, Universidade Federal do Paraná, n.º 33, 1984, http://ojs.c3sl.ufpr.br/ojs/index.php/letras/article/viewFile/19325/12610.
[8] Unger, L., Zuppa, L., «Dificultades en la producción de algunos sonidos del español por parte de estudiantes brasileños», *V Coloquio CELU: Lengua, identidad y procesos de integración regional — La evaluación y certificación en relación con estos procesos*, Córdoba, 2009, http://www.celu.edu.ar/sites/www.celu.edu.ar/files/images/stories/pdf/coloquios/5_coloquio/Dificultades_en_la_produccion_de_algunos_sonidos_del_espanol_Unger_Zuppa.pdf.
[9] Oliveira, A. M. R., *op. cit.*
[10] Refira-se neste contexto que o grupo rr- é regido por diferentes regras de translineação em ambos os idiomas.
[11] Vem aqui a propósito assinalar que a conjunção *y* se redige como *e* quando a palavra que se lhe segue começa por *i*: *guapa e inmerecida esposa; fueron e hicieron*.
[12] Ressalve-se porém que já em *cuestión* (questão) e *secuestro* (sequestro), essas propriedades fonético-fonológicas são claramente distintas.
[13] Este epíteto emprega-se em espanhol com letra minúscula.
[14] Aos grupos de três vogais em português corresponde amiúde um grupo em espanhol de que faz parte *y*: *apoyo* (mas: *mayordomo/mordomo), joya, onomatopeya, ensaya, desmayo, mayúsculo, concluyeron*.
[15] *ie* surge por conseguinte aqui em posição átona.
[16] *ie* procede do *e* latino em posição tónica.
[17] *ue* procede do *o* latino em posição tónica.
[18] Quando a conjunção *o* é seguida de uma palavra começada por *o* passa a *u*: Hacía ya diez *u* once años que nos conocíamos.
[19] Abstraímo-nos aqui das diferenças de pronúncia de vogais como *e* e *o* em português segundo a posição que ocupam na palavra, já que em espanhol não se distingue vogais abertas e fechadas.
[20] Esta diferença de vogal pode ditar em concomitância uma alteração de consoante: *porquería/porcaria*. cf. outro caso: *desempaquetar/desempacotar*.
[21] Refira-se a acentuada diferença de pronúncia dos grupos *ge-* e *gi-* em ambas as línguas (em ES como velar fricativa surda).
[22] Velar fricativa surda.
[23] Os grupos ze- e zi- não existem em espanhol: *finalisé, narice, fusilamiento*.
[24] Nesta posição, s/c são variantes ortográficas, sem implicações fonético-fonológicas em português.
[25] A realização fonética de *ch* em espanhol não coincide com a do português: *chorizo*.
[26] Em espanhol não é pronunciado como vibrante glotal ou fricativa velar, mas como alveolar vibrante múltipla sonora.
[27] Em espanhol não se distingue foneticamente *b/v*.
[28] Este fenómeno não afeta vogais tónicas, nem no caso de acréscimo nem de supressão de vogal.
[29] Inclui casos de inexistência de um equivalente sintético em português, tornando necessário o recurso a uma perífrase.
[30] Neste caso a forma do nome em espanhol coincide com a forma do particípio passado.
[31] Pelo contrário, em espanhol, as formas da primeira pessoa do plural do presente e do pretérito perfeito não se distinguem: hablamos, falamos/falámos.
[32] Para uma maior exemplificação, ver, p. ex., Neta, N.F. A., «Aprender español es fácil porque hablo portugués: Ventajas y desventajas de los brasileños para aprender español», *Cuadernos Cervantes de la Lengua Española*, época II, ano III, 2012, http://www.cuadernoscervantes.com/lc_portugues.html.
[33] Na realidade, *d* em posição final não é pronunciado.

Redação: Paulo Correia (Comissão); Valdemar Ferreira (PE); Manuel Leal (Conselho da UE); Victor Macedo (CESE-CR); António Raúl Reis (Serviço das Publicações)
Grupo de apoio: Ana Luísa Faria (Conselho da UE); Hilário Leal Fontes (Comissão); Susana Gonçalves (Comissão); Ana Lorenzo Garrido (Comissão); Joana Seixas (CESE-CR)
Paginação: Susana Gonçalves (Comissão)
Envio de correspondência: dgt-folha@ec.europa.eu

Edição impressa: oficinas gráficas do Serviço de Infraestruturas e Logística — Bruxelas (Comissão)
Edição eletrónica: sítio Web da Direção-Geral da Tradução da Comissão Europeia no portal da União Europeia — http://ec.europa.eu/translation/portuguese/magazine

«a folha» ISSN 1830-7809

ISSN 1830-780-9

9 771830 780004